Modelling Lexical Effects with Multilink: Frequency, Cognate Status, and Translation Asymmetry

Jesse Peacock

S4448456

Master's Programme In Linguistics

Radboud Universiteit Nijmegen

Supervisor: Professor Dr. Ton Dijkstra
2nd Reader: Dr. Sean Roberts
3rd Reader: Luis-Miguel Rojas-Berscia
16th Decembre, 2015

1. Table Of Contents

2. Abstract

In the present paper, Multilink (Dijkstra & Rekké, 2010) was tested, a computational model of isolated word translation that integrates theoretical notions from the Revised Hierarchical Model (RHM) (Kroll & Stewart, 1994) and the BIA+ model for bilingual word recognition (Dijkstra & Van Heuven, 2002). Simulations were conducted, using the stimulus materials from a word translation production experiment by Pruijn (2015, in collaboration with Peacock). The model's performance to the reaction times in this experiment was then compared to empirical data from Pruijn and from Christoffels, De Groot, and Kroll (2006). In these experiments, Dutch speakers of English had to translate printed Dutch or English words as accurately and quickly as possible into English or Dutch, respectively. Each input item was a high or low frequency word that could be a cognate or noncognate. Simulations of the experimental data were then analyzed through 4 sets of statistical tests: Spearman's rank correlation, analysis of variance, generalized regression modelling, and divergence testing. The simulations showed a strong cognate effect (cognates are translated faster than noncognates) and a weak frequency effect (high-frequency words are translated faster than low-frequency words). However, the simulation neither exhibited a statistically significant translation direction effect (L1→L2 translation equivalents should be translated faster than L2→L1), nor were certain experimentally-observed interactional effects. Although Multilink did produce translations with a high level of accuracy, the simulated results did not match those of the empirical data in detail. A number of adjustments and modifications of the model will be necessary to obtain better fits between model and experimental data. The findings are interpreted and compared to the predictions made by other theoretical models (RHM, BIA+). Suggestions for future experiments and model adaptations are discussed.

Keywords: computational model, simulation, bilingualism, lexical facilitation, Multilink, translation, latency, visual word naming, mental lexicon, cognate, interactive activation, lexical access, recognition & production, BIA+, RHM.

3. Acknowledgements

A thesis cannot be completed alone, and I must give credit and recognition to all the people who have helped me inestimably in this endeavour: to Dr. Ad Foolen, who has been a solid rock of optimism and compassion; to Dr. Ton Dijkstra, for teaching me how to conduct a proper psycholinguistic experiment; to the best internship and experiment partner possible, Lex Pruijn, for his dedication, hard-work, and genuine talent; to Dr. Sean Roberts, for offering to be my second reader, and for helping me finally learn some essential skills for a 21st century scientist: neural networks, Bayesian inference, and programming in Python, and R; Luis-Miguel Rojas-Berscia, for being the 3rd reader, bringing me into the world of research academia, and being a true friend; to Jakob Lesage, Jeremy Collins, Dr. Hilario De Sousa, & Tessa Yuditha at the MPI Typology Clubhouse; to Dr. Harald Hammarstrom, Dr. Steven Levinson, Dr. Russell Gray, and the entire Grambank team; to Dr. Stefan Frank, Dr. Gijs Mulder, and Dr. Pieter Muysken; to Merel Maslowski and Dr. Francisco Torreira, for also helping me learn statistics in R and analyzing results; to Dr. Pieter Seuren, for his in-depth critique of this paper, and all the amusing anecdotes and intelligent discussions; to my brother, Jason, and my niece & nephew; to my friends the world over, academic or otherwise; and lastly, to every musician who's songs have given me the courage to choose life. This thesis and none of my work, current and future, would be possible without y'all. Your wisdom has not been lost on me, and I am grateful for every second taken to help me become a better scientist and a better person.

4. Introduction

The modern era of globalization continues to bring distant populations into closer contact than ever before. Increasingly ubiquitous communication technologies and digital infrastructure have rapidly diminished the time required to exchange messages and information in the 21st century. As a consequence, multilingualism has become very important, in order to share and propogate ideas for business, culture, and science. As Bhatia and Ritchie (2013: XXI) point out, multilingualism is currently the rule throughout the world and will become increasingly more important in the future. With approximately 38 languages being spoken per country, knowledge and use of two or more languages is the common state for most communities on the planet. In all likelihood, multilingualism has been the predominant condition of human cultures and interactions since the first human diaspora from Africa.

Inevitably, when speakers of different languages meet, there will be a need for translation. Conversation is only the outward sign that translation is taking place; the act of translation must always first arise in the mind of a bilingual speaker. This begs the question: how does a bilingual speaker internally represent, recognize, and produce from these two systems? What are the cognitive implications of bilingualism? In order to answer these questions, studies have been conducted in the past to determine the nature of bilingual lexical access. The findings of these studies have been noteworthy for describing phenomena such as code-switching (Kootstra et al., 2010; 2011), language asymmetry (Meuter & Allport, 1999), language mode (Grosjean, 1998), and the neuroanatomical effects of bilingualism (Xiang, 2012).

In order to restrict the variation inherent to bilingual discourse (c.f. Muysken, 2000, 2004), most experiments have been concerned with single lexical items only. Some of the experimental tasks regularly used include: word-naming (Jared & Szucs, 2002), lexical decision (De Bot et al., 1995), picture naming (Christoffels et al., 2006), and semantic priming (Matsumoto et al., 2005). Successive experimentation has motivated theories concerning the operations of a bilingual lexicon. These theories generate predictions about task behaviour, spawning models of bilingual lexical processing & access. Some of these models, which have varying scope, specification, and architecture (covered later in Section 5), include the Dual-route Model (Coltheart et al., 1993) Revised Hierarchical Model (Kroll & Stewart, 1994), Inhibitory Control model (Green, 1986; 1998), SOPHIA (Van Heuven & Dijkstra, 2001), Bilingual Interactive Activation Plus (Dijkstra & Van Heuven, 2002), and, lastly, the MULTILINK model of Dijkstra & Rekké (2010) the last of which has been employed for the simulation in this thesis. These models of human cognitive processing are primarily built from "naming" paradigm behavioural measurements, but more recent studies have applied analagous paradigms with electrophysiological and hemodynamic measurement devices, demonstrating how bilingualism affects neural connectivity and functionality. The observations produced by these empirical data collections efforts are tested against empirically-based models, in order to refine their predictive capacities (Abutalebi & Green, 2006; Abutalebi, 2008; Van Heuven & Dijkstra, 2010).

We must realize that words, having an internal structure, do not act as wholistic units. Preceding literature shows that words in the lexicon have quantifiable dimensions (c.f. Schreuder & Weltens 1993). These dimensions comprise specific properties present at all grammatical levels: phonetic, or syllabic/phonotactic constraints; and conceptual, semantic, morphological, syntactic, or pragmatic units. When the lexicon is accessed via recognition or production of stored word forms or concepts, these dimensions engender active, experimentally-testable effects, found to be significant in lexical access. Known as "lexical effects", they are observed to occur in both monolingual & bilingual speakers in two flavours: facilitatory (aiding access), and inhibitory (impeding access). Individual words in the mental lexicon have these effects due to their interactions with other concepts, categories, words, and constructions within the lexical system. Lexical effects are tested by manipulating the aforementioned dimensions as independent variables. Manipulations correlate with observed systematic variations in dependent variables such as naming latency/response-time (Antos, 1979; Griffin & Bock, 1998), ERP components (Pylkkanen et al., 2004), BOLD (Blood Oxygen Level Dependent) signals (Edwards et al., 2005), or gaze duration (Schilling et al., 1998; Dahan et al., 2001). The observed variation informs us about the types of cognitive operations transpiring within the mental lexicon, and also within our more general Human-Language Computational System. When considering the operation of a bilingual system, these effects become more significant: a bilingual lexicon is, after all, theoretically double that of a monolingual lexicon. How do these lexical dimensions interact within the bilingual lexicon? How do lexical effects influence cognitive processing and access routes? Are languages within the mental lexicon equal, or unbalanced, and how does that affect individual and interacting lexical dimensions? What types of models have been created to explain these effects, and are their predictions accurate? Many questions arise regarding the nature of the bilingual lexicon; this study examines and attempts to generate solutions to these questions by comparatively testing the predictive performance of the Multilink model against recent empirical data, simulated in a visual bilingual word-naming task.

There are many lexical dimensions generating facilitatory and inhibitory effects in the bilingual lexicon, far too many to be described and tested in one study. For the current thesis, the following dimensions have been considered experimentally-relevant, and are categorically-manipulated by the stimulus as the independent variables (these are covered in greater detail in Section 5.2.1): **translation direction** (whether translation is from L1→L2, or

L2→L1); **frequency** (how often a particular word occurs to a speaker); and **cognate status** (the orthographic or phonological similarity between two translation equivalents). Furthermore, outside of the manipulated dimensions under direct consideration, other lexical dimensions such as concreteness, length, and onset phoneme are statistically-controlled. It should be noted that this simulation is focused on the empirical task of translation production, where a participant not only must recognize a presented token[1] as a word belonging to one language, but must also enounce the equivalent of that token in a different target language within a reasonable time-frame. In other words, translation production involves recognizing a particular input word, linking it up to its semantics and concept, and then produce an output word in another language that as approximately the same meaning. Translation recognition involves only the first step. By focusing on production rather than recognition, the results of this simulation are more applicable to the modelling of natural bilingual contexts than recognition experiments. Limitations to this methodology are detailed in section 6.4.

This thesis is broadly-structured as follows: Section 5 surveys the literature of bilingual lexical access; section 6 specifies the hypotheses, and the methodology used to create and evaluate the data; section 7 covers data analysis, and the results of the statistical tests employed; section 8 interprets the tests, discusses and compares the results to the findings of other studies, and proposes directions for future research; section 9 concludes the paper with a general summary; cited works and supplementary materials are found in sections 10 and 11, respectively.

[1] "token" being defined as any individual item in the stimulus, not as a logical or linguistic type-token distinction (c.f. Wetzel, 2009)

5. Background Review

This section covers the literature and concepts pertinent to the current study, focusing on 4 topics: section 5.1 reviews various current theoretical and computational models of bilingual lexical access; section 5.2 covers several of the dimensions relevant to the bilingual lexicon, dividing these into 3 subsections: independent variables, controlled variables, and uncontrolled nuisance variables; section 5.3 examines the incongruences within the bilingual lexicon, the phenomenon known as "bilingual asymmetry"; and lastly, section 5.4 discusses 3 prior bilingual visual word-naming studies that have informed the current debate, and contributed to the motivation for the current experiment; section 5.5 concludes and summarizes the review.

5.1 **Models Of The Bilingual Lexicon**

Some models of bilingual lexicon are discussed, particularly the BIA+ and Multilink computational models.

5.1.1. **Revised Hierarchical Model (RHM) (Kroll & Stewart, 1994; Kroll et al., 2010)**

The seminal paper of bilingual lexical access and processing is Kroll and Stewart (1994), which introduces the Revised Hierarchical Model (RHM), following the results of 3 experiments. From these results, an asymmetry was observed in participants: "Subjects were consistently faster to translate into the first language than into the second language." (Kroll & Stewart, 1994: 157). This pre-computational model accounts for the translation asymmetry by showing two potential routes for translation (see Figure 1, next page): the lexical association route, in which the L2 is translated via the L1; and the conceptual route, where a lexical item is directly linked with its concept. In particular, this explains why cognates are translated faster. Summing up their findings, Kroll and Stewart (1994: 168) state, "The data we have presented support the claim that translation from the first language to the second is conceptually mediated, whereas the translation from the second language to the first is lexically mediated. Taken together, the data support the predictions of a revised model of bilingual memory representation in which cross-language connections between lexical representations and concepts are asymmetric." Iterating, the latency of L1→L2 production is less than the latency of L2→L1 production, due to the extra step required by lexically mediated bilingual production.

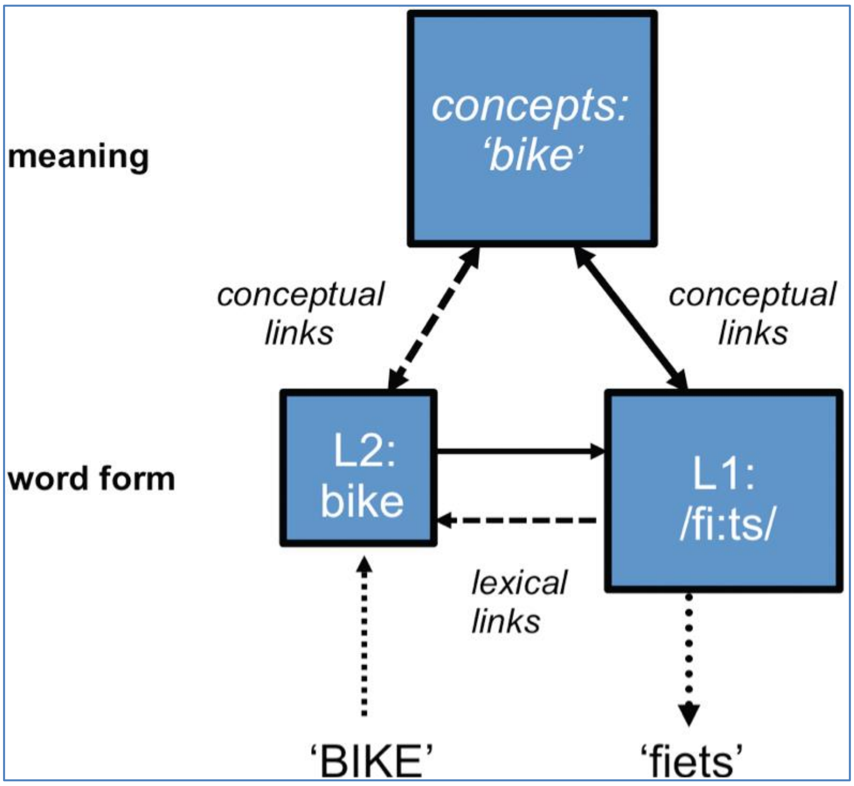


Figure 1. The Revised Hierarchical Model (Dijkstra & Rekké, 2010: 403)

The RHM is not without problems, as noted by Brysbaert and Duyck (2010). Per the summary found on page 368: "There is little evidence for separate lexicons, and for language selective access; excitatory connections between lexical equivalents impede word recognition; the L2 conceptual route is stronger than proposed by the RHM; and it appears necessary to distinguish language-dependent and language-independent semantic features." Rebutting to Brysbaert and Duyck (2010), Kroll et al. (2010) agrees that the RHM is in need of revision after 10+ years of citation and testing, but charges that it was never intended to be a model of bilingual visual word recognition, but rather a model of late-in-life L2 acquisition, production, and imbalanced lexical proficiency. Arguments concerning the RHMs predictions about language nonselective access, translation, conceptual and semantic access routes, and L2 development are then brought up, and compared to current studies.

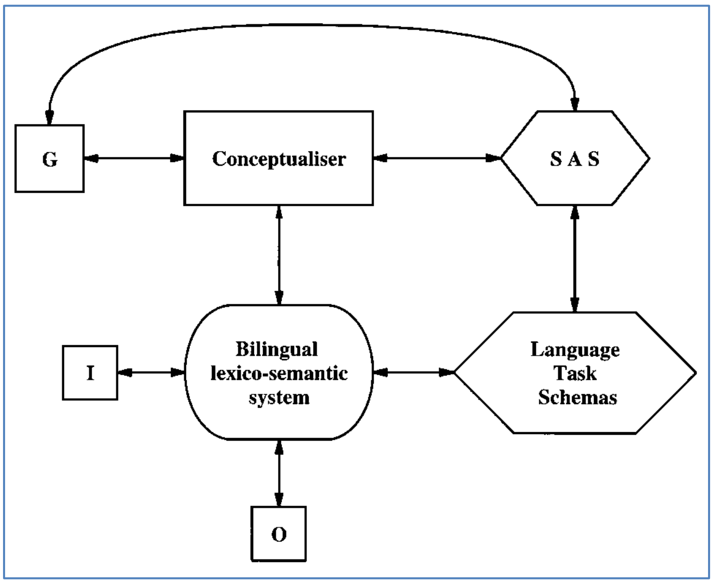


Figure 2. Architecture of the Inhibitory Control model (Green, 1998: 4)

5.1.2 **Inhibitory Control model (IC) (Green, 1986; 1998)**

The IC model of Green (1986, 1998) is a pre-computational descriptive framework for explaining speech production errors, and how neurotypical and impaired or aphasic bilinguals control two languages, focusing on bilingual lexico-semantic access and selection systems. It was deliberately designed to accommodate data from both neurotypical and pathological studies within the model. To explain how the switch from L1 to L2 (and vice versa) is accomplished, the concept of "language nodes" is employed, which identify the language membership (and activate the node for each language) of inputs and outputs. This design is in opposition to the concept of "language mode" (Grosjean, 1998), but allows a shared lexicon to be implemented (also in opposition to the language-specific lexicons employed by the RHM). The primary concepts of the model are summed up as: control, activation, and energy (originally called "resource"; "The resource idea makes explicit the fact that a system needs energy to operate" (Green, 1986: 215)). The model itself separates into 3 parts (see Figure 2, previous page): control of language task schemas, lemma-level lexical selection, and inhibitory control. Ultimately, it is the intended to predict bilingual performance and selection, mediated by a limited pool of resources, much like the cognitive process of coordinating, planning, and producing other physical actions. Many aspects of its design are shared by other models, such as the BIA+ (Dijkstra & Van Heuven, 2002).

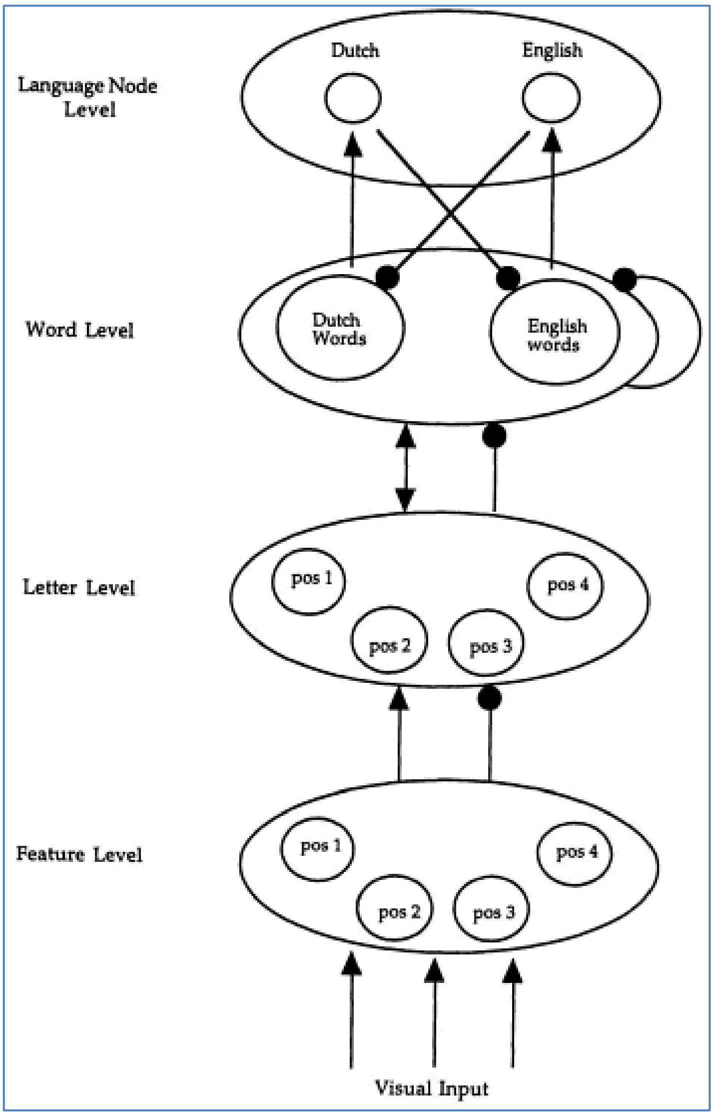


Figure 3. The BIA Model (Dijkstra & Van Heuven, 1998: 200)

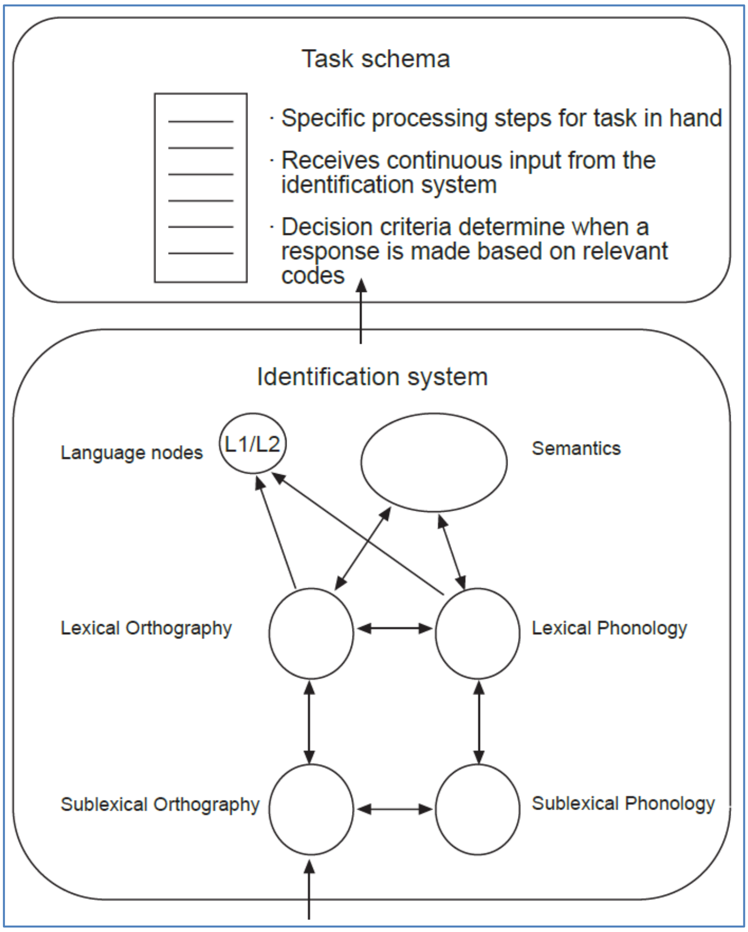
5.1.3 **Bilingual Interactive Activation Plus (BIA+) (Dijkstra & Van Heuven, 1998, 2002)**

The BIA+ model began as an extension of McClelland & Rumelhart's Interactive Activation (IA) model (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982), a visual perception model of symbol and word recognition. This class of model was originally defined as follows:

"[. . .] information processing takes place through the excitatory and inhibitory interactions among a large number of processing elements called units. Each unit is a very simple processing device. It stands for a hypothesis about the input being processed. The activation of a unit is monotonically related to the strength of the hypothesis for which the unit stands. Constraints among hypotheses are represented by connections. Units which are mutually consistent are mutually excitatory, and units that are mutually inconsistent are mutually inhibitory [. . .] When the activation of a unit exceeds some threshold activation value, it begins to influence the activation of other units via its outgoing connections; the strength of these signals depends on the degree of the sender's activation. The state of the system at a given point in time represents the current status of the various possible hypotheses about the input; information processing amounts to the evolution of that state, over time [. . .] This 'interactive activation' process allows each hypothesis both to constrain and be constrained by other mutually consistent or inconsistent hypotheses."

(McClelland & Elman, 1986: 2)

Figure 4. Architecture Of The BIA+ model (Dijkstra & Van Heuven, 2002: 182)



Using an approach known as "nested modelling", the BIA model — *sans plus sign* — was created (Dijkstra & Van Heuven, 1998), extending the utility of the IA from monolingual word recognition, into the bilingual domain. Like the IC, it employs languages nodes; but unlike the RHM and IC models, the BIA is a functional computational recognition model. It uses a 4-layer architecture (see Figure 3, previous page), each layer corresponding to a different resolution level within the lexical access system: letter features

(14 for each letter position), letters (26 for each position within the word), words (complete with a combined lexicon of 1,324 English and 978 Dutch words), languages (one node per language), and contains separate excitatory (arrows) and inhibitory (dot-heads) connections; the direction indicates the flow of activation. As Figure 3 illustrates, activation is directed from bottom-up, beginning with the identification of features, then letters, to words (stored in the lexicon file), with language as the last activated nodes, isolating the language membership for each word. This design makes the model capable of simulating a variety of task specifications, according to a language non-selective access model.

The BIA model had its limitations, primarily lacking full specifications: a lack of integrated phonological or semantic representations, underspecified representations for form-similar tokens, a lack of "participant"-specified task descriptions, and the relationship between token identification and the task is not suitably specified, among a few others. Dijkstra and Van Heuven (2002) presented an updated architecture for the model. Because it incorporated a large portion of the original model with the same nested-modelling method, it was dubbed the BIA+ model (see Figure 4, previous page), and solved the forementioned limitations while also adding in a Task Schema system, inspired by the "language task schema" subroutine for the IC model. Formally, the model separates the two systems into the Word Identification system and the Task Schema system, the former feeding information about active representations into the subroutines of the latter. A major modification was effected within the Word Identification system: no longer do the language nodes asymmetrically inhibit the word nodes from top-down; this function was transferred to the Task Schema component. These changes and improvements helped the model to better predict and resolve questions about the inner workings of the bilingual lexicon. However, with respect to word translation, the BIA+ model was lacking in another aspect entirely: performing accurate bilingual recognition is only half the equation. In order to comprehend the core of the translation process through computational modelling, the other half is needed as well: production.

5.1.4 **Multilink (Dijkstra & Rekké, 2010; Dijkstra et al., in prep.)**

Recently, a new model has been designed to implement the word translation process as a whole: Multilink (Dijkstra & Rekké, 2010; Dijkstra et al., in prep.). As a successor to the BIA+, Multilink is part of the latest generation of computational bilingual lexical processing models, incorporating developments from previous generations into a localist-connectionist[2] design that simulates common tasks and scenarios regularly found in psycholinguistic

[2] Also known as an "artificial neural network"; the term "connectionist model" is generally preferred by (psycho)linguists because they are simplified computational representations of neurons, only partly modelling neuron behaviour and action (c.f. Grainger & Jacobs, 1998; Christiansen & Chater, 2001).

behavioural studies: lexical decision, language decision, cognate recognition, semantic spreading activation, and word translation. Like its predecessor, it is (largely) an interactive-activation-based model, and constructed through the same principle of nested-modelling. Multilink is intended to faithfully simulate each step of the process: recognition, meaning retrieval, and word production, in both beginning and proficient bilinguals. It shares similarities with the RHM, IC, BIA+, and WEAVER++[3] models: it correlates resting-level activation of each input item with its word form frequency; distinguishes orthographic, phonological, semantic, and language membership representations, which form the integrated lexico-semantic system; incorporates a task & decision system; L1 and L2 word form and conceptual representation links are flexible and the model assumes that the lexicons can vary in size; and can test the presence of word association links between languages, and also the presence of inhibitory links between word forms.

Inputs for Multilink proceed in a similar fashion to the BIA+ model as well: a token (present in the lexicon, of course) activates lexical-orthographic representations, and examines tokens in the lexicon — regardless of language membership, the final activated representation level in the model — for their form-similarity (using a length-normalized Levenshtein Distance algorithm for cognate selection), and word form frequency. When an orthographic representation for a token becomes active, semantic and lexical-phonological representations are also activated, following the same procedure as with the orthographic representations: activation is input to each semantic node, and, in turn, activation is received to the activated token from each semantic node, proportional to its "association strength" (0-1 scale) contained within a database file. For each time-step — called a "cycle" — the level of activation is calculated as the sum of the activation — both excitation and inhibition — in the previous time-step, plus the input from each active connected representation. When a candidate's activation surpasses a specific threshold, the model is ready to select an output; the lexical-phonological component — representing word production — activates, and the final output is generated, representing spoken production in the target language. Candidates with higher frequency and lower LD scores are more likely to be chosen as the principal output; being constructed around these parameters, activation results in several candidates, contingent upon the summed activation that ranks each active representation. The fundamentals of the activation processes employed by Multilink remain largely the same as in McClelland & Rumelhart's original IA model.

[3] WEAVER++ is a predominantly monolingual computational model, constructed to demonstrate how natural language lemmas are planned, controlled, and produced for spoken utterances (Roelofs, 1992, 1997). Roelofs (2003) applied it to bilingual utterances for the first time.

Figure 5. Activation process within Multilink, showing the translation of the English token "fork" into the Dutch token "vork". (De Wit, 2014: 9)

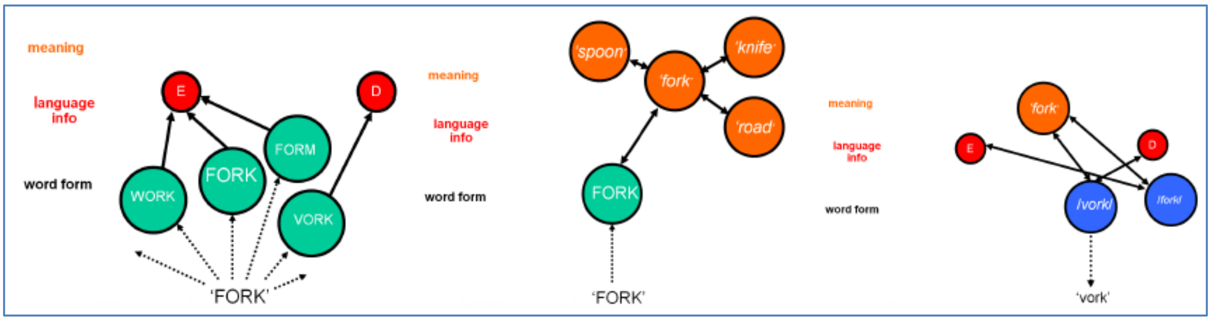


Figure 5 (above) illustrates the human word translation process as a whole. It shows how the English word "fork" is translated into the Dutch word "vork": first, the input letter-string activates all other form-similar tokens, the orthographic neighbourhood, regardless of language, including the orthographic representation of the input ("fork" → FORK); next, the orthographic representation activates the semantic representations (FORK → /fork/); the active semantic representation triggers the target language lexical-phonological token (/fork/ → /vork/), which outputs the lexical-orthographic form (/vork/ → "vork"). While this architecture is certainly more complex than the BIA+ model, Multilink extends the utility of its predecessor into roles where the BIA+ could not adequately perform. Producing more accurate results, utilizing an empirically-based design, doubling of the length restriction criterion from 4 to 8 letters, a wider scope-of-use, and recognition & production mechanisms all make Multilink a more advanced model for the simulation of bilingual lexical cognitive processing.

Multilink is (currently) programmed in Javascript, and natively contains a lexicon of approximately 1,000 Dutch-English word-pairs. Word form frequencies are derived from CELEX (Baayen, Piepenbrock, & Van Rijn, 1993). As of December, 2015, it is version 1.02, with a possible major revision planned for the near future. For the formulas used to calculate normalized Levenshtein Distance and resting-level activation, see the appendix.

5.2 **Lexical Dimensions**

As stated in the introduction, interactions between individual and sets of lexical items — and respective dimensions — facilitate or inhibit access to word forms stored in the mental lexicon. These interactions are especially important in bilingual systems, because each word form has, in theory, twice as many other word forms to interact with, since there are two languages through which activation can propagate (if the lexicons are considered to be integrated, like in Language Non-selective Access Models (French & Jacquet, 2004; De Bot, 2004). This section lists several of the best-studied lexical dimensions, and details their effects (with the exception of translation direction and proficiency, which are covered in section 5.3). It is partitioned into 3 groups: 5.2.1 details the manipulated dimensions, 5.2.2 details the controlled variables, and 5.2.3 discusses two important, but uncontrolled, lexical effects.

5.2.1 **Independent Variables**

The dimensions subsumed under this section represent the manipulated dimensions of the simulation. Frequency and cognate status are discussed here.

5.2.1.1 **Frequency**

Frequency of usage is an important factor for determining a word's speed of lexical access. Numerous studies, stretching back to Zipf (1935, 1949) have demonstrated its importance, and have demonstrated significant correlations between frequency and other linguistically-salient dimensions (which themselves can engender other facilitation effects): length (Piantadosi, Tily, & Gibson, 2010), gaze duration (Rayner, 1998; Pollatsek et al., 2008) , particle detection (Kapatsinski & Radicke, 2007), sentence length (Sigurd et al., 2004), and even speech rate (Lorenz, 2015).

Within translation production, frequency is highly correlated with latency, as measured in visual word-naming studies like Pruijn (2015), Christoffels et al. (2006), De Groot et al. (1994), inter alia. High-frequency tokens have significantly lower latency, while Low-frequency tokens have significantly higher latency. Data on word form frequency is typically obtained from corpora, although this can be problematic for bilingual word-naming: frequency of usage requires exposure, which will be subjective for each participant, and is highly-correlated with L2 proficiency. It is more accurate to state that corpus frequency data represents a *potential frequency* that each person is exposed to, and will subsequently produce.

Frequency data for Dutch and English stimuli in the present simulations were originally obtained from the SUBTLEX-US (Brysbaert & New, 2009) and SUBTLEX-NL (Keuleers, Brysbaert, & New 2010) databases, as was the binary frequency distinction (please refer to section 4.1.1.ii): tokens with a 10-Log frequency of 1.50 or lower were classed as "Low Frequency", and tokens with a 10-Log frequency of 1.60 or higher were classed as "High Frequency".

Stimuli used in the simulation are balanced for frequency on a 10-Log scale, $\bar{X}_{\text{English}}$ = 1.59, and $\bar{X}_{\text{Dutch}}$ = 1.55.

5.2.1.2 **Cognate Status**

Cognates (as defined according to Dijkstra & Rekké, 2010; Schepens, Dijkstra, & Grootjen, 2012; and Pruijn, 2015) are words that have form-similar translations in both L1 and L2. Form similarity is determined using the Levenshtein Distance (henceforth, "LD"), an "edit-distance" algorithm that quantifies the difference between two letter-symbol sequences (Levenshtein, 1965; 1966), using three distinct "edit operations": **insertion** — add a symbol into the sequence; **deletion** — remove a symbol from the sequence; and **substitution**[4] — exchange one symbol in the sequence for another. For every operation conducted upon a single symbol, the measurement score increases by 1, the total representing the number of edits necessary to make one sequence the same as another. Although originally designed for correcting errors in binary signals, the LD algorithm eventually found its way into linguistics as a technique for measuring similarity ratings for cognates, such as Kessler (1995), which used the LD of transcribed phonetic strings to compute linguistic distance for dialect groupings in Irish Gaelic.

LD allows the manipulation of cognate similarity within the lexical processing system, resulting in a "cognate facilitation effect" relative to non-cognates. Unlike interlingual homographs — which have an orthographic LD of 0 (as with the Dutch and English word "film") and shared meaning — cognates have an orthographic LD of approximately 1-3, or share 75% form-similarity; consider, for instance, the Dutch word "tomaat", and the English word "tomato", which have an LD of 2 (English→Dutch: deletion, insertion). Due to their interlingual nature, being tagged by both L1 and L2 language nodes in the lexicon, these words have significantly faster access, as demonstrated by studies like Christoffels et al. (2006). On the opposite end, consider the non-cognate pair, "Art-Kunst", which has no form-overlap at all. Non-cognates are neither inhibited nor facilitated within the lexicon.

Like the above studies, the current simulation only considers LD in orthography and visual bilingual word-naming, but there are studies that have considered phonological LD (Gooskens & Heeringa, 2004; Nerbonne & Heeringa, 1997). Tokens are balanced according to this dimension: 50% of the experimental stimuli are cognates, and the other 50% are non-cognates. For an overview of cognate facilitation and processing, the reader is referred to Dijkstra (2005).

5.2.2 **Controlled Variables**

The dimensions in this section represent the controlled non-manipulated dimensions of this study. Previous studies (see section 5.4) have proven the necessity of rigorous

[4] Substitution is sometimes seen as a combination of the insertion and deletion operations; other edit distance algorithms can have fewer, or greater, numbers of edit operations built into the metric.

statistical balancing in order to detect lexical effects with high significance. The variables detailed here include: length, concreteness, and phonetic onset.

5.2.2.1 **Length**

Length — the number of graphemes or phonemes in a word — has a measureable effect on lexical retrieval. As discovered by New et al. (2006) through statistical investigations of the English Lexicon Project lexical decision data, word length has a facilitatory effect for lengths of 3-5 letters, no significant facilitatory or inhibitory effect for lengths of 5-8 letters, and an inhibitory effect for words with lengths of 8-13 letters. Additionally, orthographic frequency, number of syllables, and orthographic neighbourhoods all had their own inhibitory and excitatory effects. The so-called "Word Length Effect" — first described by Baddeley et al. (1975) — is the observation that shorter words have a higher recall rate than longer words, but recent investigations have since called this effect into question (Neath et al., 2003; Lovatt et al., 2000), and others have found no significant effect attributed to length at all (Bachoud-Levi et al., 1998). In general, it is still believed that length plays an important role in lexical processing and access, the reasoning being that shorter words require less articulatory planning than longer words, thus being produced at a faster rate. At least one study has reported a "sign length effect" for signers, analogous to the word length effect in speech (Wilson & Emmorey, 1998), demonstrating the potential for a general symbol-sequence length effect despite studies showing otherwise.

Regardless of any present-or-otherwise effect, stimuli employed for this study are length-balanced per language: all tokens are 3-8 letters in length, $\bar{X}_{\text{English}}$ = 5.02, and $\bar{X}_{\text{Dutch}}$ = 4.91.

5.2.2.2 **Concreteness**

"Concreteness" is a subjective measure of how substantive or abstract a word is, defined by Gee, Nelson, and Krawczyk (1999: 1) as: "[. . .] the extent to which one can readily form a mental image of a word's referent, and it is measured by asking subjects to rate words on a numerical scale." It is closely related to, and highly correlated with, yet also distinct from, the phenomenon of "imageability" (Richardson, 1975, 1976; Altarriba et al., 1999). Consider the pairs "saxophone" and "essentialness"; both are nouns, however one is much more easily pictured than the other. According to the raw data from Brysbaert et al. (2014), these words lie at opposite ends of the concreteness spectrum, rating at 5, and 1.04, respectively. This quantitative and qualitative difference foments the aptly named "concreteness effect", in which highly concrete words like "saxophone" are processed faster — and by a different route — than abstract words like "essentialness" (Kroll & Merves, 1986). Other experiments, however, have questioned the nature of this effect, testing whether there is truly a cognitive separation between concrete and abstract words in the mental lexicon (Van Hell & De Groot, 1998). Follow-up studies by Barber et al. (2013) and Jessen et al. (2000) have shown electrophysiological and hemodynamic evidence for divergent cognitive routes between words based on concreteness.

Dutch and English stimuli are balanced for concreteness, with ratings on a 1-5 scale for Dutch and English tokens taken from Brysbaert et al. (2014), $\bar{X}_{\text{English}}$ = 4.21, and $\bar{X}_{\text{Dutch}}$ = 4.04.

5.2.2.3 **Phonemic Onset Type**

An important determinant of word-naming latency is the type of phoneme a word begins with. This is true even within visual word-naming experiments, which lack enunciated\acoustic input for the participant, because the mental lexicon maps visual wordforms to acoustic signals via Grapheme-Phoneme Correspondence (GPC) rules (Bassetti, 2013; Tham et al., 2005). Onsets specifically are the most pertinent segment of the word form in naming tasks because they occupy the first slot of initial syllable of the word; Gow et al. (1996) proposed that onsets have singularly salient perceptual properties that drive lexical segmentation, recognition, and access. In addition, the triggering of a voice key depends on phoneme onset type, such as voiced vs. unvoiced, or fricative vs. plosive consonant. Rastle et al. (2005) and Palo et al. (2015) noted significant delays for some types of consonants, particularly voiceless fricatives, showing a divide between acoustic naming latency and articulatory naming latency[5]. This effect was even present when measuring with ultrasound imaging, and allowing participants to "mentally-prepare" by pre-exposure to stimuli (thus bypassing the word/utterance planning stage of production).

When visually presented words are named, another issue is how the phonological representation is derived from the orthographic input. In other words, we must understand how in word naming orthographic input representations are mapped onto phonological output representations. Davelaar et al. (1978) ran several experiments targeting homophones to find out how grapheme-phoneme encoding operates. The authors propose a dual-route "race" model, in which graphemic and phonemic forms are activated on the basis of a visual input letter-string, and then race to activate the correct response. Scheerer (1986) proposes a more cooperative dual-route approach, with direct and indirect routes into the

[5] The difference between "articulatory" and "acoustic" naming latency is the time between the placement of articulators (labia, uvula, etc.) in order to commence phonation, and the time required for air to be pushed up from the lungs, pass through the vocal folds and the articulators, and exit the mouth\nose as phonation.

mental lexicon via orthography. Frost (1995) proposes that the phonological representations associated with orthographic input units might be impoverished relative to the phonological representations used in spoken word recognition. Following a series of naming tasks using unpointed Hebrew script, he developed an interactive dual-route for generating phonological representations from orthography. The results support a strong phonological hypothesis: phonological units for a word form are computed from orthography individually or in clusters, and assembled as a final product, rather than retrieving complete phonological structures based on whole orthographic word forms. An fMRI study by Fiebach et al. (2002) used a lexical decision task, contrasting neural activity elicited by pseudowords and low & high frequency words, to corroborate the hypotheses of dual-route access. Finally, as discussed and further modelled by Feustel et al. (1983), the visual word recognition process leaves a trace in episodic memory, creating a repetition effect, a confound that lowers the processing time required for repeated lexical segments and features.

In the end, a trinary division for onsets was made for the current study, coding for either voiced or voiceless consonants, or vowels: 49% of tokens have voiced consonantal onsets, 44% have voiceless consonantal onsets, and the remaining 6% have vocalic onsets. Although this is not a perfect remedy to the articulation or measurement issue, nor does it fully negate the repetition effect, it was reasoned to be the most viable and expedient solution. Because of the delay in detecting certain initial phonemes — with a phoneme-induced bias being as large as 100 ms (Kessler et al., 2002) — onset categorization was also useful to diminish the bias introduced by these known technical difficulties with voice-key latency measurements.

5.2.3 **Uncontrolled Variables**

The following three dimensions — conceptual-association (i.e., spreading activation), morphological families, and orthographic neighbourhoods — have not been statistically-controlled by the current study, but were deemed significant enough within the literature to discuss. These shall be considered "nuisance" variables.

5.2.3.1 **Conceptual Association**

Effects of conceptual association, also known as "spreading activation", are omnipresent in natural language and general human cognition. Spreading activation involves two words that are semantically-related, the so-called "prime" and "target"; take the following pairs as examples: "Nurse-Hospital", "Soldier-Tank", or "Kitchen-Oven". The first word interacts with the second, activating it through a cognitive link that can be considered frequency-modulated (the two words might frequently co-occur): nurses often work in hospitals, soldiers are accompanied by tanks, and kitchens usually contain at least one oven. One way that activation spreads within the mental lexicon is through these categorical associations, and their connected word forms. Compare to semantically unrelated (or, at the very least, much more distantly related) pairs: "Eagle-House", "Car-Ocean", or "Countryside-Milkshake". When experimentally-tested, a semantic priming task is often used. The study of Meyer and Schvaneveldt (1971) is considered one of the early significant studies concerning conceptual association. It involves two lexical decision experiments, the results of which support a theory of facilitatory activation between meaning-related words. For an overview of studies that have investigated this phenomenon, see Neely (1991).

Conceptual association is not directly accounted for in the stimuli, as this was postulated to diminish the number of available tokens by a large factor. An associative effect in the latency measurements was avoided by breaking up meaning-related pairs in the pseudo-randomized lists before presentation. Conceptual association between separately, but closely-presented, input and output word forms was also not addressed, but this is considered a minimal noise factor. Multilink itself contains a structure to handle semantic relations, a file that indexes related pairs with a strength-rating[6]. This association list is only available for tokens native to Multilink's lexicon, however, and stimulus materials are not represented. The current study does not address this aspect, and does not consider it to be a major confounding variable.

5.2.3.2 **Morphological Families**

The morphological productivity of a word form can affect lexical access in significant ways, with active individual tokens initiating related word forms within the lexicon. Mulder (2013), citing Schreuder and Baayen (1997), defines "morphological family" as: "[. . .] the number of complex words that are morphologically related to a given word and in which this word occurs as a constituent." (Mulder, 2013: 16). Using the examples of "home" and "villa", Mulder observes that some words have greater potential for compounding than others. A study by Schreuder and Baayen (1997) paints a much more complex picture of the frequency effect within morphological families, particularly for monomorphemic word forms: the monomorphemic word form frequency alone ("home") combines with the frequency of morphologically-related complex word forms ("home-s", "home-ward", "home-base", "home-ly", etc), creating a peculiar effect within the mental lexicon; the result causes especially frequent complex word forms to split off from their root monomorpheme, gaining their own lexicosemantic representations. As the studies of Mulder (2013) establish, this morphological

[6] For instance, one of the first listed relations is between the English tokens "aardvark", and "dictionary", with a strength of .013

family effect extends across languages. For bilingual speakers, not only the input word and its target translation equivalent can be activated, but also all members of the morphological family — in both input and target languages — when form-similarity is close enough. This is due to the fact that the root morpheme is party to all members of the family. When processing interlingual homographs (such as "normal" in Dutch and English), family activation is expected. Dijkstra et al. (2005) investigated the effect of morphological family size on bilingual word recognition in 3 experiments, finding that both L1 and L2 family sizes affect lexical processing, presenting task-dependent facilitatory and inhibitory effects, even when performing in their native language. Although this effect aligns overall with the word frequency effect, even after accounting for it, the morphological family effect remains significant. Still, it is worth noting that both effects may have a similar origin. Lehtonen and Laine (2003) attribute this to representation vs composition: highly frequent complex word forms gain their own full representation in the lexicon, in the interest of efficiency, whereas less frequent complex word forms must be decomposed. This is especially evident when comparing low and high frequency affixed or compound words: "bakelite" and "hydrocarbon", low-frequency, must be decomposed, but "skydiving" and "cheesecake" are specialized, represented lexemes, thanks to their high frequency-of-usage. But their findings are somewhat contrary, showing that bilinguals more often take the decomposition route, which might be attributed to the fact that bilinguals receive less lexical input for either language.

While there is at least one measurement available for morphological family size, called the "Information Residual" (Del Prado Martin, Kostic, & Baayen, 2004), it has not been employed in this study, and does not appear widely utilized at this time.

5.2.3.3 **Orthographic & Phonological Neighbours**

Orthographic neighbours, as defined by Mulder (2013: 20), "[. . .] are words that differ from each other in only one letter position [. . .]. The English word *wool* only differs in one letter position from other English words such as *fool, wood,* and *tool*. A similar orthographic neighbourhood relationship can exist across languages." Neighbourhoods — i.e. word form fields — are either orthographic, phonological, or both. The two are intimately related (Frost, 1998), and direct links between orthographic neighbourhood density and phonological neighbourhood density have been observed (Grainger et al., 2007). The general finding is that words with denser neighbourhoods are processed faster than words with sparse neighbourhoods, given that lexical activation can spread faster through denser neighbourhoods.

Measurement methods for estimating neighourhood size exist. Coltheart's N (Coltheart et al., 1977) — now considered somewhat defunct — and the OLD20

(Orthographic Levenshtein Distance 20 [closest orthographic neighbours] of Yarkoni, Balota, & Yap, 2008) sample orthographic neighbourhood size. Data using both of these measurements is available in both CELEX (Baayen, Piepenbrock, & Van Rijn, 1993), and SUBTLEX (Brysbaert & New, 2009; Keuleers, Brysbaert, & New, 2010) databases, but have not been employed for the current study. As we shall see, Multilink uses an activation-spreading design that begins by transmitting activation throughout the input token's neighbours (see section 5.1.4); this strategy is reasoned to make balancing for neighbourhood size redundant.

5.3 **Bilingual Proficiency & Translation Asymmetry**

Many multilingual speakers will exhibit an asymmetry in the strength of their acquired languages. Even very fluent mutlilinguals may be unbalanced. Despite being able to quickly and efficiently select and switch between languages, a processing asymmetry is often observable in experimental tasks. Models like the RHM and Multilink are constructed with the intention of understanding the myriad of factors that correlate with and determine the degree of language asymmetry. Many factors co-determine the degree of this asymmetry. Global factors — in the sense that they are non-applicable to any single item or sets within the lexicon — include: age & manner of acquisition (Sabourin et al., 2014), proficiency (Christoffels et al., 2006), and language dominance (Heredia, 1995, 1997); local factors would include, but are not limited to, the lexical dimensions outlined in Section 5.2, some of which will covary with the global factors. And much like the interactions between various lexical dimensions, combinations of global and local factors can interact in significant, and partly predictable, ways. For instance, an interaction between proficiency and language dominance was discovered by Costa and Santesteban (2004): there was a language switch cost in a picture naming experiment when participants were asked to switch between L1 and L2, but this same cost — an observed increase in naming latency — was not seen in highly-proficient bilinguals. This explains why specific lexical categories like (non-)cognates and low/high frequency word forms have detectable effects in the bilingual lexical processing system: the dominant language is experienced (encountered and used) more often by the speaker, and as a consequence, each word in this language has a higher subjective frequency. It becomes more easily accessed, whereas the opposite case is noted for the non-dominant language(s). Kroll and Stewart (1994) explain this language-based access dissimilarity through the hypotheses of "conceptual association" and "word association", the core of the RHM; Sholl et al. (1995) corroborates these hypotheses through picture-naming and translation tasks, finding that the translation task can be primed by the picture-naming experiment, in forward direction. A similar study by Meuter and Allport (1999) found comparable, corroborating results. They explain the switch cost as the consequence of an active suppression of the dominant language. Because the dominant language is so strong, avoiding it requires a stronger inhibition when an item of the non-dominant language is recognized, and as a consequence it must be re-activated from "further down" by an input word from the dominant language (the "inertia hypothesis"). These findings are in line with Language Non-Selective Access Models (NSAM), which regard languages in the mental lexicon as integrated, rather than separate.

Logically, connections between languages have at least two directions[7], asymmetric or otherwise: the forward translation direction (L1→L2), and the backward translation direction (L2→L1). For over 25 years now, psycholinguistic studies have slowly uncovered and pieced together the basic mechanisms and interactions of global and local factors at work in the mental lexicon. Nevertheless, the nature of language asymmetry has thus far remained elusive and somewhat contentious, centering on the directional effect of language asymmetry: is there a facilitatory effect for forward, or backward translation, and to what extent does it operate? How is it changed by higher or lower levels of language proficiency and dominance? Pruijn (2015) highlights 3 studies — Kroll & Stewart (1994), Christoffels et al. (2006), and De Groot et al. (1994) — that have offered major contributions to this debate (see Table 1, below).

Table 1. Latency (in milliseconds) from three previous studies. Percentage of correct responses in parentheses (Pruijn, 2015: 13)

	Forward Translation	**Backward Translation**	**Overall found effect**
Kroll & Stewart (1994) Experiment 3 Participants: Students	1269 (51.3%)	1140 (63.5%)	Backward facilitation
Christoffels et al. (2006) Experiment 1 Participants: Students	912 (90.4%)	978 (89.0%)	Forward facilitation
De Groot et al. (1994) Experiment 1 Participants: Students	1307 (80.4%)	1271 (86.3%)	Null

5.4 **Recent Empirical Studies**

A number of studies from the past 15 years, all concerning translation production experiments, are reviewed in this section. All of these studies manipulated the same variables as discussed in section 5.2. In particular, the studies of Christoffels et al. (2006),

[7] Assuming a purely bilingual system

and Pruijn (2015) are elaborated. For a larger in-depth review of multilingual lexical access and processing literature, the reader is referred to Szubko-Sitarek (2015), *Multilingual Lexical Recognition In The Mental Lexicon Of Third Language Users.*

5.4.1 **Kroll, Michael, Tokowicz, & Dufour (2002)**

The study by Kroll et al. (2002) consists of two experiments, each comparing two samples of adult native English-speaking participants, investigating L2 lexical acquisition and the transition from word association to conceptual association inside the mental lexicon as L2 fluency increases. The stated goal of the study was "[. . .] to examine the process of lexical access for both L1 and L2 during second language acquisition." (Kroll et al., 2002: 141).

The first experiment involved two English-French sample groups — a low-proficiency group, and a high-proficiency group — performing two tasks: first, a visual word-naming task (a word is presented to the participant on a screen, and enunciates the word aloud); and second, a visual word translation task (an L_N word is presented to the participant on a screen, and enunciates the translation equivalent in the target language), measuring latency & accuracy. These tasks are associated with lexeme-level processing, the performance of each participant indicating the route of access to lexical information. Results of this experiment supported the views of the RHM (Kroll et al., 2002: 153), asymmetry being greater for the less fluent, accuracy and latency measurements supporting backwards facilitation, and the observation that the low-proficiency group relies more on form-relation between languages than the high-proficiency group.

The second experiment, very similarly designed to the first, tested a new set of participants in two sample groups: a low-proficiency condition, and a high-proficiency condition. Unlike the previous experiment, the fluency difference was greater between conditions, with the low-proficiency condition being described as "[. . .] nonfluent learners at the very early stages of L2 learning [. . .]" (Kroll et al., 2002: 153). Furthermore, the groups were not learners of a single language, but were divided between Spanish (the majority for both groups) and French. These groups performed the same word-naming and translation tasks as the first experiment, with two additions: a reading-span task[8], and a lexical decision task[9]. The results showed an effect of proficiency for both word-naming and translation tasks in both groups, but the highly-proficient bilinguals only showed evidence for language

[8] This is a task in the "memory span" paradigm. Participants read sentences, and are asked to recall the final words; the measurement is based on how many final words can be recalled. Span tasks are often used to assess short-term memory, and cognitive ability or intelligence.

[9] In a lexical decision task, a participant is shown a letter string, and presses one button if it is a word in the target language, and another button if it is not. Results can be used as a metric for language fluency.

asymmetry in word-naming. Language asymmetry, favouring the L1, in the translation task, was observed in both proficiency groups. Comparing the forward and backward translation conditions, L2-learners had a 111 ms difference, while the bilinguals had only a 48 ms difference.

The results of Kroll et al. (2002) demonstrate a clear backwards facilitation effect, substantiating claims made by the RHM that both high & low-proficiency L2 processing is accomplished through lexical association with the L1, rather than directly attaching lexemes to concepts; nevertheless, the experiments also show that, as fluency increases, L2-conceptual connections will form. This is corroboration of the conclusion of Lehtonen & Laine (2003): increasing frequency (and thus, proficiency) causes L2 lexemes to gain their own representation, rather than requiring these inputs to be processed through existing L1-structured cognitive pathways.

5.4.2 **Christoffels, De Groot, & Kroll (2006)**

Table 2. Results Of Experiment 1 & 2, mean RT (in milliseconds (Christoffels et al., 2006: 333)

Group	Language				Language effect
	English–Dutch		Dutch–English		
	RT	% error	RT	% error	RT
Students (n = 39)					
Cognates					
HF	784	1.3	712	1.1	
LF	944	8.7	904	3.7	
Noncognates					
HF	959	5.0	883	3.8	
LF	1225	29.1	1149	29.9	
Average	978	11.0	912	9.6	66
Interpreters (n = 13)					
Cognates					
HF	656	1.3	656	.43	
LF	753	0.9	745	.43	
Noncognates					
HF	803	2.1	782	1.3	
LF	1055	5.1	1083	12.8	
Average	817	2.35	817	3.74	0
Teachers (n = 15)					
Cognates					
HF	657	0.4	663	0.0	
LF	738	1.5	721	1.1	
Noncognates					
HF	763	0.7	769	0.0	
LF	992	4.1	986	12.6	
Average	787	1.7	785	3.4	2

Centering on the role of simultaneous interpreters, such as those present at international conferences and symposia, Christoffels et al. (2006) presents two experiments in this paper, aiming to explain the cognitive skills necessary for this group to comprehend and produce in two languages at the same time. The concurrent input of one language and output in another implies incredibly rapid planning and action, including the following steps: recognition of an input as a member of one language, proceeding to conceptual activation, the L2 equivalent word-form becomes active, the L2 word-form is produced, positioned into an L2 phrase\clause (that is, itself, approximate to the original phrase in the meaning conveyed), and finally, is physically articulated. This requires concentrated, coordinated effort, inevitably requisitioning resources from various areas of the brain. "The goal of the present study is to begin to understand how basic components of language processing may be different when an individual is a skilled interpreter and how simultaneous interpreting is related to individual differences in memory capacity." (Christoffels et al., 2006: 326).

The first experiment tested two groups of native Dutch speakers, all of whom had high-proficiency in English: a 1st group of simultaneous interpreters, and a 2nd group of university students, using picture-naming[10], and single-word translation production tasks as the primary measures of language performance. Since there are also questions about whether simultaneous interpreters have larger-than-average memory capacities — either through a selection bias, or simply from accumulated experience as a bilingual — a series of memory assessments were also used: a word span task, reading span task, and a speaking span task. Lastly, two control tasks were included: vocabulary test, and an arrow reaction time test[11]. All of these tasks were completed in both languages, as functional capacity can differ between languages (Chincotta & Underwood, 1998). The following lexical dimensions were manipulated: word-form frequency, and cognate status. The authors predict: "If the subskills examined here are indeed important to simultaneous interpreting, we predicted that the interpreters would outperform the students on both measures of language processing and memory capacity. On the control measures, we expected that the interpreters would have better vocabulary knowledge than the students, but that performance on the basic reaction time test should be unrelated to interpreting skill." (Christoffels et al., 2006: 327). The results of experiment 1 strongly favour the interpreter group — who performed notably well on the memory assessment tasks — over the student group. Moreover, the interpreter group did not show a facilitation effect for language direction in the translation task, unlike the student group, which translated L1→L2 faster than L2→L1, a forwards facilitation effect (see Table 2, previous page).

The second experiment, similar in design to the first experiment, tested two high-proficiency English groups of native Dutch speakers: simultaneous interpreters, again, and trained English teachers, all tested using the same set of tasks. Teachers were selected for their supposed similarities to the interpreter group, both groups hypothesized as being approximately equal in their global interindividual factors: proficiency, language dominance, age, bilingual working experience, education, and general interest in languages. Although students, as a group, are known to be proficient — but unbalanced — bilingual speakers they do not share these same characteristics which make trained bilingual teachers a particularly good match for a comparative study with the interpreter group. The results support these conjectures: the teacher group performed similarly to the interpreter groups on

[10] A participant looks at a picture of an object, and produces the word that the picture represents in the target language.
[11] A participant views left or right-facing arrows on a screen, and presses a button corresponding to the direction of the arrows.

the picture-naming and translation tasks, however, the interpreter group still outperformed the teacher group on the memory capacity tasks, establishing the suspected working memory bias thought to operate within simultaneous interpreters (however the cause is still inconclusive).

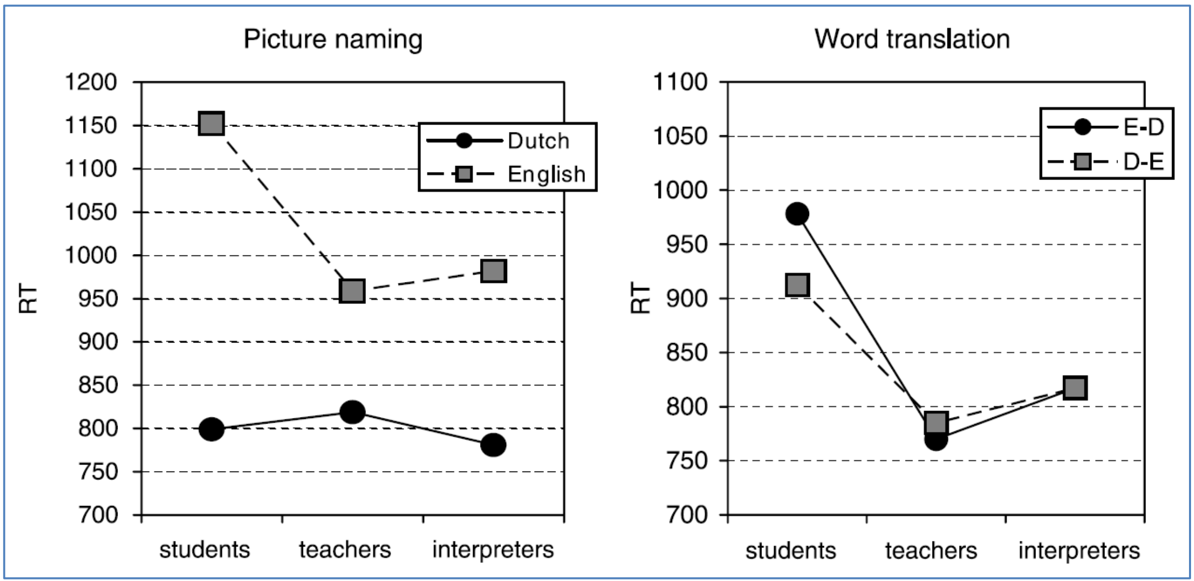


Figure 6. Average RT (in milliseconds) for each group, in picture-naming and word translation tasks (Christoffels et al., 2006: 340)

Figure 6 (above) provides visualization to the general summary: notably obvious is the fact that the students performed worse in both tasks — except for L1 picture-naming — and the teacher and interpreter groups seem to pattern closely. We can be confident that proficiency, dominance, and frequency are important variables in these tasks, as prior research has proven time and again. However, aside from the memory capacity advantage, employment as a simultaneous interpreter does not enhance one's basic lexical processing operations any more than other professions which require multilingual proficiency. Additionally, cognate status and frequency were shown to have separate effects. A cognate effect was even observed in the picture-naming tasks, lending greater support to Language Non-selective Access Models (NSAM), but also to the concept of an orthographic-phonological → visual activation route via semantic activation.

This study, in opposition to Kroll & Stewart (1994), and Kroll et al. (2002) presents evidence for forward facilitation. However, the study is not without some problems, as noted in Pruijn (2015). The stimuli employed by Christoffels, et al (2006) are not as well-balanced as the study might have them seem, with some interlingual homographs, semantically-related pairs, and no accounting for onset type. This is problematic, and a confound of concerning proportions, particularly given the small sample sizes of the teacher and interpreter groups. The results only add to the contentious nature of the debate concerning language facilitation.

5.4.4 **Pruijn (2015)**

		English					Dutch				
		N	**Freq**	**RT**	**Conc.**	**Len**	**N**	**Freq**	**RT**	**Conc.**	**Len**
Cognates	**HF**	32	2.23	523	4.27	4.94	32	2.18	538	3.92	4.72
	LF	31	.96	555	4.05	5.06	32	.94	565	4.24	4.81
Non-cognates	**HF**	32	2.11	529	4.17	4.78	32	2.11	547	3.84	5.22
	LF	30	1.04	546	4.35	5.33	32	.96	563	4.17	4.87

Table 3. Stimuli characteristics; mean ratings for frequency, naming latency, concreteness, and length shown for all 8 stimulus categories (Pruijn, 2015: 21)

Another study in a long line to investigate bilingual lexical access with hopes to proffer solution, Pruijn (2015, in collaboration with Peacock) details a single experiment designed to test language asymmetry, the outcomes of the RHM, and lexical facilitation. Data from 42 native Dutch-speaking participants were collected (24 females and 18 males) for a visual single-word translation production task, taking place in two conditions: the forward translation condition (Dutch → English), and the backwards translation condition (English → Dutch). Participants had their response latencies measured from the presentation of stimulus until the detection threshold[12] is triggered by enunciation. Stimuli — 256 in total — varied in the same 3 conditions as previous studies: frequency (low or high), cognate status (cognate or noncognate), and translation direction (forwards or backwards), and were additionally balanced for other known lexical dimensions, to avoid the same confounds that plagued the stimulus set of Christoffels et al. (2006) (see Table 3, above). This information was gathered from the SUBTLEX (Brysbaert & New, 2009; Keuleers, Brysbaert, & New 2010) database. Stimulus, with few exceptions, was the same for both forward and backward translation conditions; e.g. a participant received reversed translation-equivalent pairs (Forward would be "koffie", to "coffee", whereas backward would be "coffee", to "koffie"). Each block of the task presented 128 tokens, and upon completion, the participant filled out a language history questionnaire, obtaining standard demographic information, and subjective ordinal ratings about profiency and general foreign language experience.

[12] set at .07 in Presentation

Final analysis retained the same data-cleaning procedures as Christoffels et al., disregarding inaccurate responses. Outlier data points were removed: latencies below 350 ms were classed as technical or participant errors, and latencies above 2000 ms were classed as null responses. 5 participants were omitted for exceeding the inaccuracy threshold of 10%, and 3 specific tokens were removed due to low total accuracy ratings[13]. Ultimately, 1,856 (17.26%) data points were eliminated, leaving 8,896 total data points for repeated-measures and univariate ANOVA testing. Included participants were separated into low and high proficiency bilingual groups based on questionnaire responses, with approximately half in each group.

Table 4. Mean latency and accuracy percentages, with overall language effect (Pruijn, 2015: 24)

		English-Dutch	Dutch-English	Language Effect
Cognates	**HF**	737 (98.6%)	726 (97.0%)	-11
	LF	826 (95.2%)	808 (93.1%)	-18
Non-cognates	**HF**	847 (96.9%)	800 (95.9%)	-47
	LF	981 (88.7%)	917 (90.2%)	-64
Average		848	813	-35*

From the results (see Table 4, above), a significant interaction between cognate status and translation direction was found: cognates were more resistant to language asymmetry than noncognates, the latter of which were observed to have a forward facilitation effect. Word form frequency, however, was not found to interact with translation direction; low and high frequency tokens were observed to have similar levels of measured language asymmetry. Likewise, a combined interaction between the 3 conditions was not observed. Cognate status and frequency, however, did have a significant interaction: the frequency effect was observed to have higher correlation with noncognates than with cognates. As with previous studies, L2 proficiency modulated the overall latency observed, but the respective facilitation effects were observed to function independently of proficiency, no significant interaction being observed.

Across all categories, the results point towards forward translation facilitation, rejecting the predictions of backward translation facilitation from the RHM, and also contesting the results of previous studies that have confirmed it as a valid model of the bilingual lexicon. While the RHM equally predicts cognate status as a factor in language asymmetry, it does not predict why L1→L2 noncognates are translated faster than L2→L1

[13] These same 3 tokens are excluded from the current analysis as well, see section 6.4

noncognates. One possible, culturally-bound explanation is proposed by Pruijn, remarking on the power of the internet and other instantaneous communication infrastructures since both the studies of Kroll & Stewart (1994), and Christoffels et al. (2006): "Another possible reason [. . .] is that they [the studies of Pruijn, and Christoffels et al.] are conducted within a smaller time range [. . .] Twelve years might not seem much, but the rise of the internet has certainly had a great impact on bilingual development in Dutch children and students, and this might have very well influenced the way they process English words. In other words, general Dutch-English bilingual proficiency might have accelerated the past years, which brings along a difference in participant proficiency and, arguably, this affects translation mechanisms. "(Pruijn, 2015: 28). At the same time that these results disregard the RHM, they do form a pattern of predictions much closer to the BIA+, and, correspondingly, with Multilink as well (as shown by simulations conducted by Dijkstra & Rekké (2010)).

5.5 **Conclusions Based On The Literature**

The previous section has assessed the fundamental topics necessary to comprehend the current state of research in the domain of bilingual lexical access and processing. To reiterate, the RHM, IC, BIA+, and Multilink models were appraised and detailed, with specific reference to the architectures utilized by each; pertinent facilitation effects, stemming from lexically-interactive dimensions such as frequency, orthographic neighbourhoods, and conceptual-meaning, were explained; the general phenomenon of translation asymmetry was described; and, lastly, recent studies, such as Christoffels et al. (2006) were checked, which will ground the results presented in the following sections.

Making predictions from the results of previous studies, projections about the working of the bilingual lexicon are formed: dependent variables, such as latency or gaze duration are modulated by lexical dimensions; when a word is recognized as grammatical input, activation distributes throughout the associated lexical networks, regardless of the input or target languages, and facilitates or inhibits potential candidates until the correct word form is selected and produced; highly-proficient bilinguals are less susceptible to language directional asymmetry — frequency of language practice is one of the more important dimensions for individual speakers; lexical-associative effects are major and omnipresent noise factors, found at all levels of word processing; dimensional interactions within the cognitive system, such as those between cognate status and frequency, are potentially significant; and, above all, the need to control for nuisance variables (other lexical dimensions that are only contributing noise to the results) and balance all global and local factors within an experiment — cannot be understated. Ultimately, the evidence reviewed has given us a window into the process of visual single-word translation production, and multilingual lexical cognitive processing. Application and attention to this knowledge is imperative to the creation of empirically-matching computational models.

6. Simulation Methodology

As we have seen in the previous sections, certain theoretical and methodological considerations must be attended to when designing a simulation. The methods chosen and outlined below follow the development of this approach. Section 6.1 discusses the specific research questions being answered, section 6.2 defines the stimulus according to the independent and dependent variables, section 6.3 declares a number of limitations of the methods employed, and section 6.4 clarifies the manner in which errors & inaccuracies are treated.

6.1 **Research Questions**

This study controls stimuli in a similar fashion as Christoffels et al. (2006), dividing them by translation direction, frequency, and cognate status in order to simulate the results of Pruijn (2015) and correlate the empirical data with the simulation data from the current study. The primary research question asks if Multilink can sufficiently simulate:

i. The translation direction facilitation effect — the difference between forward and backward translation, and is forward, or backward translation facilitated?

ii. The frequency facilitation effect?

iii. The cognate facilitation effect?

In essence, the model simulations should approximate the results of the real experiment; this is the model-to-data distance, related to the concept of "goodness-of-fit". With this in mind, the expected distributional difference between the two datasets — the latency results of Pruijn (2015) and the Multilink cycle-times — should be relatively small. Of course, in order to make a real comparison between the empirical and model data, it is necessary to bring the model data into more-direct correspondence. With this intention, two statistical models have been implemented: a linear model, and a z-score model. These models translate the Multilink cycle-times into predicted-latency outputs by scaling them according to the factor necessary to reach the empirical latency data. Statistical analysis of the simulations will examine results calculated via 4 tests: a Spearman's rank correlation coefficient test, F2 ANOVA, generalized regression modelling, and divergence statistics. These tests will assess the model-to-data fit.

This simulation should demonstrate that the Multilink computational model can simulate the translation production process, by mirroring the results observed in the experimental translation task. Demonstrable observations and correlations should include: the cognate, frequency, and translation direction facilitation effects, all comparable in strength to the empirical effects; the latencies for the respective conditions should be highly predictive of cycle-time and the two will be strongly correlated. For instance, cognate conditions should show strong correlations with LD; high-frequency conditions should show a strong correlation with frequency; and interactional effects should be observable and similar in simulations and empirical data.

6.2 **Stimulus Materials**

The stimulus list used in the reported simulation comprises 256 tokens[14] in total. This list represents the most recent and thoroughly-prepared set of dimensionally-balanced materiel currently available, exemplifying the adjustments and attention necessary to generate accurate data to be used in computational modelling (see section 5.4.4, Table 3 for mean per-category ratings of 10-log word form frequency, concreteness, naming latency, and length). The independent and dependent variables measured are outlined below, each signifying one of the manipulated lexical dimensions, and their subsequent facilitation effects (as outlined in section 5.2).

6.2.1 **Independent Variables**

i. **Translation direction** — defined as: Forward translation (Dutch→English), or Backward translation (English→Dutch). 128 tokens are Dutch→English pairs, and the other 128 are English→Dutch pairs. Both translation directions use largely the same stimuli, with a few exceptions.

ii. **Frequency** — defined, somewhat arbitrarily, as: high-frequency (SUBTLEX word-form 10-Log frequency above 1.60), or low-frequency (SUBTLEX word-form 10-Log frequency below 1.50). It should be noted, SUBTLEX (Brysbaert & New, 2009; Keuleers, Brysbaert, & New 2010) was employed to generate this definition by Pruijn (2015), as it was considered to be a better measure of frequency than previous collections, such as CELEX or Kucera & Francis (n.d.); the current simulation, nevertheless, must keep in line with the frequencies already in the lexicon of Multilink, which employs the word form frequency measurements from the 1993 CELEX (Baayen, Piepenbrock, & Van Rijn, 1993) database. Thus, while the stimulus retains the same frequency division as the original experiment — using SUBTLEX — the actual frequencies come from CELEX.

iii. **Cognate status, i.e. form-similarity —** defined as: cognate, or non-cognate. This definition is based on the LD of individual tokens (see section 5.2.1.2). Tokens requiring 1-3 edits (using a length-normalized score) are considered cognates; tokens requiring 4-8 edits (no stimuli are above 8 letters in length) are considered non-cognates. Tokens with an LD of 0, interlingual homographs, such as "hand" or "film" — found in the stimulus selection of

[14] It is important to note that each token is actually a pair of prime & target translation equivalents.

Christoffels et al. (2006) — have been excluded and replaced. Multilink employs the standard 3 edit-operations within the length-normalized LD algorithm.

This results in 8 combinations of manipulated independent variables, with 32 tokens per category, as shown in the table below:

Forward				**Backward**			
High-Frequency		Low-Frequency		High-Frequency		Low-Frequency	
Cognate	Noncognate	Cognate	Noncognate	Cognate	Noncognate	Cognate	Noncognate

Forward High-frequency Cognate	Forward High-frequency Noncognate	Forward Low-frequency Cognate	Forward Low-frequency Noncognate
Backward High-frequency Cognate	Backward High-frequency Noncognate	Backward Low-frequency Cognate	Backward High-frequency Cognate

Table 5. Manipulated independent variable combinations.

6.2.2 **Dependent Variables**

There are two dependent variables to this study: latency, and cycle-time

i. **Latency** is the mean response time after visual presentation of stimulus, in milliseconds. It is taken from the appendix of results, found in Pruijn (2015: 37-42).

ii. **Cycle time** and its derivations are the model's equivalent of latency. When Multilink runs a simulated translation production task, it outputs the most faithful translation-equivalent candidate, and states the processing time — defined as "time cycles" — required to calculate this candidate. This data has been collected for each token in the list. To make predicted-latency, cycle-time has been calibrated to a millisecond scale, using cycle-time and empirical latency as paired input for two statistical models: a linear model, and a Z-score model.

As a generalization, it is expected that the "high-frequency cognate" categories will show the lowest mean latency and cycle-times, reflecting faster processing; on the other hand, "low-frequency noncognate" categories are expected to show the highest mean latency and cycle-times. Consequently, it is inferred that the brain processes tokens of these latter categories in a different fashion than the former categories.

6.3 **Design Limitations**

There are some limitations to the current simulation that needs to be declared before viewing the test results:

i. Latency data is a mean aggregation per token from each of the 37 participants, taken from the appendix of Pruijn (2015). This was done to simplify the calculations and to better match the latency data with Multilink's output. Analogous data sets should have comparable properties; where latency data has large variance between participants, Multilink's output, if run 37 times, would have little-to-no observable variance. This reduction-to-the-mean will affect the following analyses, reducing the statistical power, but the general trends exhibited by the data should remain the same.

ii. In general, frequency measurements have a corpus-dependency problem: words with an occurrence of 17 times in corpora of 1 million words and 20 million words are not directly comparable. It is therefore convenient to use the size of the corpus to weight the frequency and present it as Occurrences Per Million tokens (OPM): a token with 1 OPM in a 1 million word corpus will have an OPM of 0.1 in a 10 million corpus.As an alternative, Van Heuven et al. (2014) has introduced the Zipf scale, a logarithmic scale measurement of word form frequency, that categorizes tokens on a scale from 1 Zipf (very low) to 6/7 Zipfs[15] (very high). 1 Zipf is equal to a word form frequency of 1-per-100 million. The Zipf scale has not been employed by Pruijn (2015), nor the current study. Future analyses will need to consider the advantages of using the Zipf scale over the more common 10-log OPM measurements.

iii. Tokens are largely restricted to the nominal class, with a few adjectives. Bultena (2013) has examined cognate facilitation in verbs, obtaining varying results, ultimately pointing to a much weaker cognate facilitation effect in verbs than are currently observed for nouns. The syntactic class of the word forms is restricted for experimental control.

iv. Word-naming experiments, while statistically reliable and experimentally valid when properly designed, are concerned with isolated orthographic and phonological words; this is not necessarily a true reflection of neither lexical access, nor bilingual speech, and the reality of lexical effects — both facilitatory and inhibitory — is much more complex when viewed in-use (Swinney, 1979; Simpson et al., 1989; Duyck et al., 2007). Context plays a critical role. Speech requires a lexicon, but the two should not be conflated. In order to

[15] Only function words like definite articles are stated as having 7 Zipfs. Common nominals might have a ceiling of 6 Zipfs.

investigate natural language lexical access properly, an approach with greater scope will be required, utilizing linguistic units larger than single words.

6.4 **Errors & Removal Procedure**

Out of 256 tokens, 7 (2.73% of total data points) were removed from the analysis, for two reasons: 3 items (1.17%) had been excluded from the empirical analysis (Pruijn, 2015: 37-42); and 4 (1.56%) produced inaccurate translations from Multilink, translating "bot" as "boat" instead of "bone", for example (see Table 6 below).

The inaccurate translations all arise from the lack of active semantic representation for the input stimulus, and the position of LD in the output ranking. The inaccurate outputs are all misconstrued as cognate with other tokens used in the simulation, not randomly with other tokens within the lexicon, which do contain semantic associations listed in a separate file (see section 5.1.4 for a summary of Multilink and its architecture). These semantic links have varying numbers and strengths of semantic connections to other lexicon items, and appear to be the final measure for selecting the correct output.

All 4 inaccuracies are confined to the L1→L2 condition, while the 3 excluded tokens are confined to the L2→L1 condition.

Table 6. Inaccurate Multilink translations, and removed stimulus.

Stimuli	Translation	Latency	Cycle-time	Frequency	Direction	Cognate status	Frequency Category	Accuracy (1 = correct)	Multilink output
Inaccurate									
arend	eagle	991	33.35	4	Forward	Noncognate	Low Frequency	0	friend
beker	cup	1081	25.94	15	Forward	Noncognate	Low Frequency	0	baker
boer	farmer	968	25.54	45	Forward	Noncognate	Low Frequency	0	beer
bot	bone	1105	37.94	6	Forward	Cognate	Low Frequency	0	boat
Removed									
curl	krul	0	27.49	1	Backward	Cognate	Low Frequency	1	
eagle	arend	0	27.45	7	Backward	Noncognate	Low Frequency	1	
lack	gebrek	0	26.74	110	Backward	Noncognate	Low Frequency	1	

7 Analysis & Results

Tests explored in this section have a single purpose: to examine the goodness-of-fit by comparing the model data to the empirical data. Cycle-times were transformed into 2 sets of predicted-latency measures, each using a different type of statistical model. Due to the high total accuracy score, and the facts that it is limited to the forward translation direction condition and potentially a confound caused by the lack of semantic representation, accuracy data is not included in the analyses. First, section 7.1 briefly describes the statistical models used to extrapolate predicted-latencies from Multilink's cycle-time output; section 7.2 inspects the results through a simple visual comparison-of-means; section 7.2 discusses the correlational analyses; section 7.3 discusses the results of ANOVA analyses; next, section 7.4 shows the results of a generalized regression model; and, finally, section 7.5 states the results of a battery of model-to-empirical statistical-distance and divergence metrics.

7.1 **Cycle-time Scaling Method**

To facilitate the model-to-empirical data comparison, it was deemed necessary to scale the cycle-time outputs into predicted-latency measurements. Two statistical models were chosen to accomplish this task:a linear model, and a Z-score model. Section 7.1.1 discusses the relationship between cycle-time and latency; section 7.1.2 discusses linear model scaling; and section 7.1.2 discusses the Z-score model scaling. In later sections, the output of the latter two will be further compared to assess the better scaling method.

7.1.1 **The Relationship Of Cycle-time To Latency**

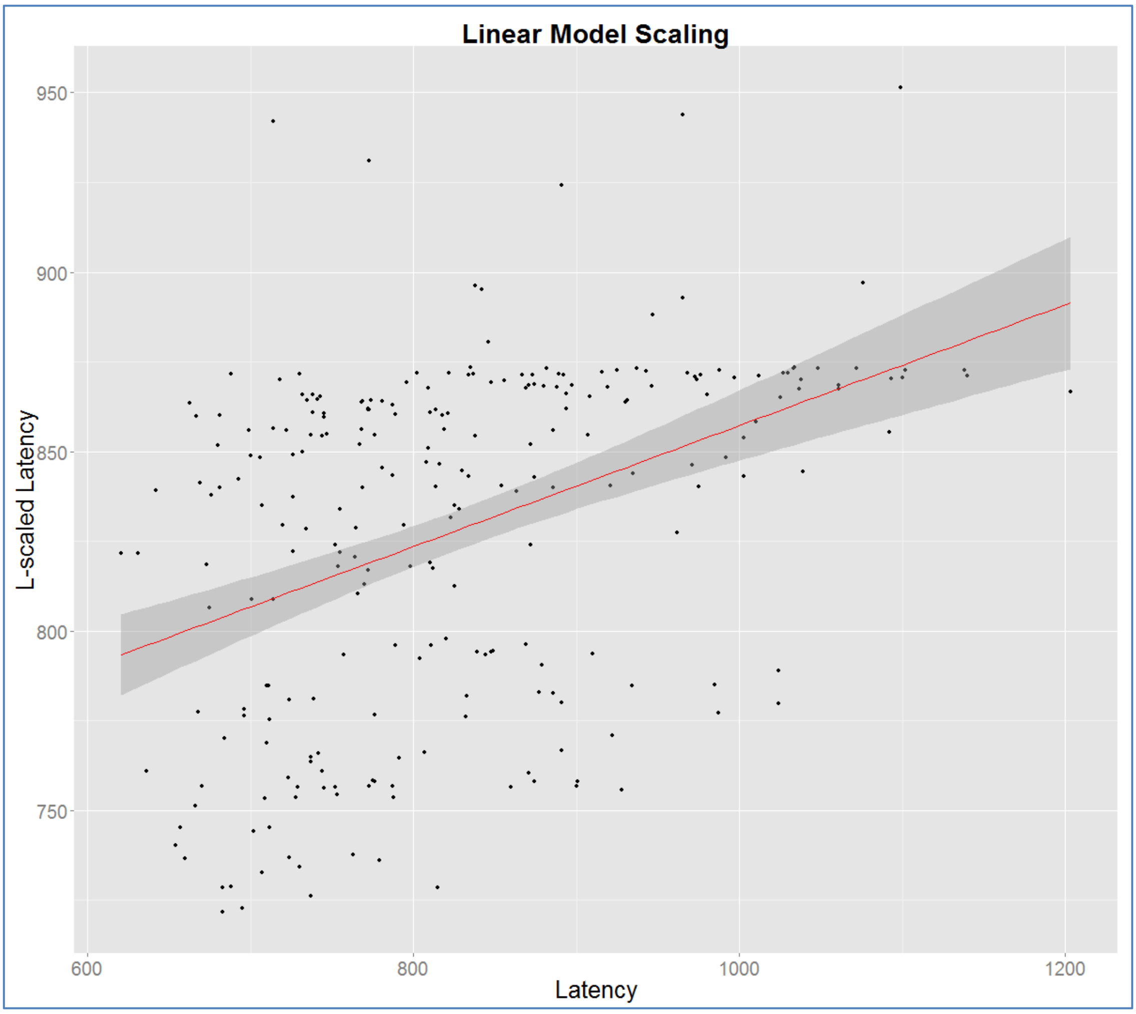
When cycle-time and latency are entered into a statistical model, scaling data points linearly, their relationship is more easily discerned. Using latency as the independent variable, and cycle-time as the dependent variable, a significant[16] correlation is observed ($F(1, 247) = 50.05$, $p = < 0.0001$), with an *adjusted* $R^2 = 0.17$. Cycle-time increases by 1 for every 24 ms ($SE = 3.38$, $p = < 0.0001$) of latency, from a base estimate of 218 ms ($SE = 86.62$, $p = 0.013$). Therefore, a latency of 1000 ms ≈ 27 cycles; conversely, 25 cycles ≈ 750 ms.

[16] Using $\alpha = 0.05$

7.1.2 **Linear Model**

The first model is a simple general[17] linear model, represented visually in Figure 7 (below). This is, essentially, the most basic and convenient scaling method available to equivocate one data group with another. In this thesis, the linear model predicted-latency measurements will collectively be referred to as "L-scaled [latency]". According to the graph, the linear model does not seem to be the most faithful reproduction of the latency measurements: 700 ms latency ≈ 800 ms L-scaled; 800 ms latency ≈ 850 ms L-scaled; 1000 ms latency ≈ 860 ms L-scaled, and 1200 ms latency ≈ 880 ms L-scaled.

Figure 7. Scatterplot of the linear model scaling, x-axis = latency, y-axis = L-scaled Latency, linear regression line in red. The shaded area around the regression line represents the standard error.



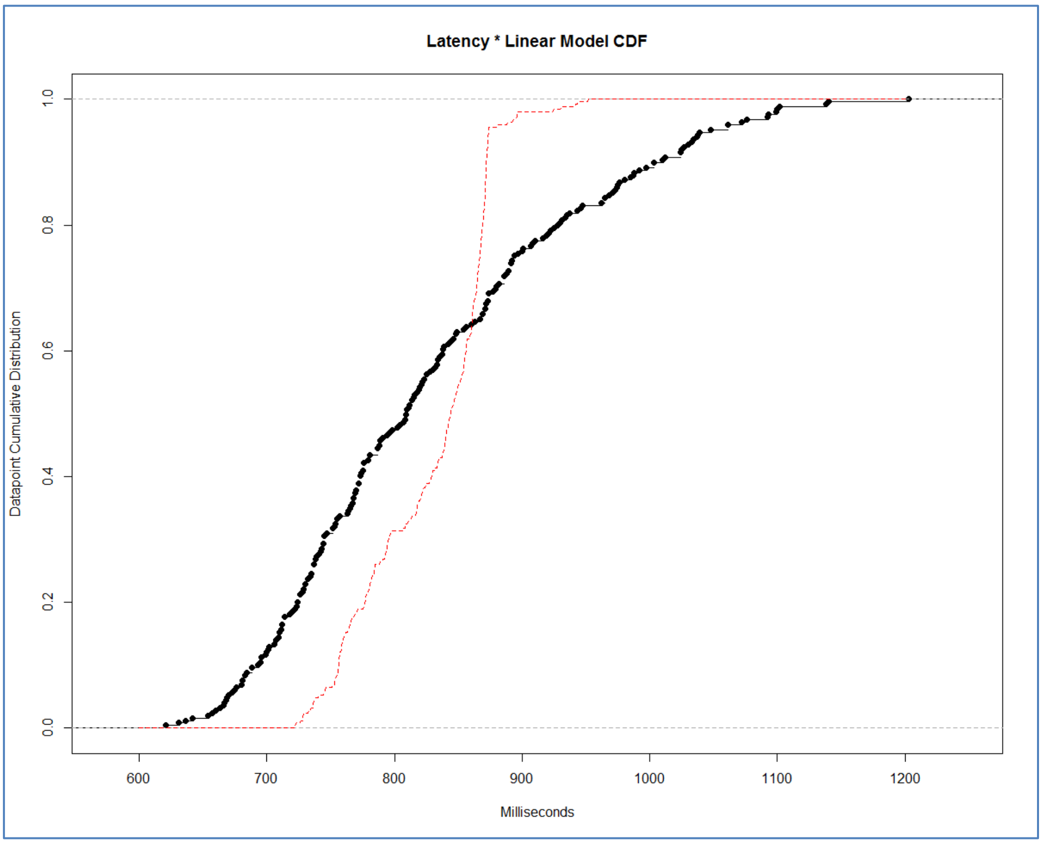
Displayed as a pair of cumulative distribution function (CDF) lines, latency and the L-scaled latency are seen in Figure 8 (next page). It is clear that these two do not match, with a high degree of visible divergence between the two CDFs. The minimum L-scaled latency begins at ~720 ms, almost ~100 ms after the minimum of latency, at a point where ~20% of latency data is already measured. After a ≤ 60 ms width of divergence, the two lines

[17] Not *generalized*, which typically refers to a type of regression model.

intersect at the ~60% mark, denoting ~850 ms. A second, larger divergent section is observed, at which point the L-scaled latency hits the maximum of its range at ~900 ms, where ~80% of latency datapoints are measured, ~300 ms before the latency CDF hits its respective maximum.

For the scaled simulation data, it appears that the tails of the latency CDF have been trimmed, shrinking the total range of the L-scaled latency by nearly 400 ms, effectively decreasing the measurements by 40% (20% eliminated from both bottom, and top) compared to the empirical data.

Figure 8. Cumulative distribution function of latency (black dots), and linear model predicted latency (red dashed-line). Each dot/dash segment represents a single datapoint.



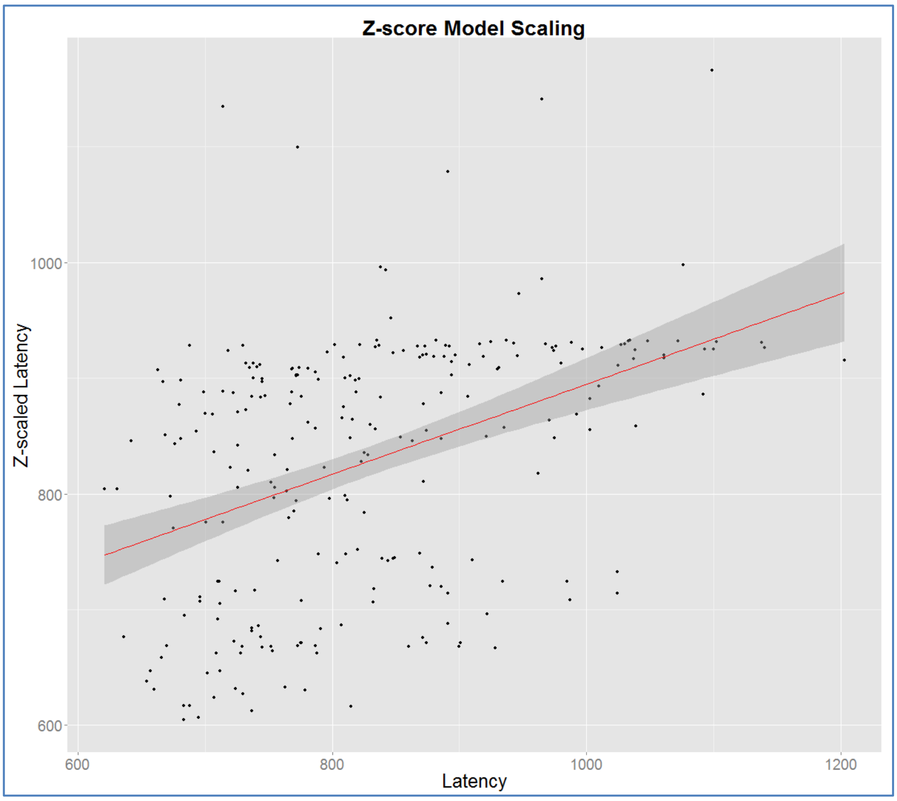
It is, admittedly, difficult to be confident about the potential performance of a scaling method that is observed to reduce the measured range by such a large margin, while simultaneously overestimating the lower quartiles and underestimating the higher quartiles.

7.1.3 **Z-score Model**

The Z-score model is a reasonably more complicated scaling method, requiring multiple steps, versus the two steps required for the linear model. It is called a "Z-score model" because it relies on calculating the Z-values of the distribution formed by the latency measurements: first, the Z-values and 10-log values are calculated for each individual latency measurement; second, similar Z-values are obtained for the 10-log values of latency, and (non-logarithmic) Cycle-time. Predicted-10-log latency measurements are attained by multiplying the cycle-time Z-values by the standard deviation of the 10-log latency, then adding the mean of the 10-log latency to each new predicted-10-log latency measurement. Finally, the scaled predicted-latency measurements are obtained by performing an antilog operation on the predicted-10-log latency data. For the duration of this article, the Z-score model predicted-latency measurements will collectively be called "Z-scaled [latency]".

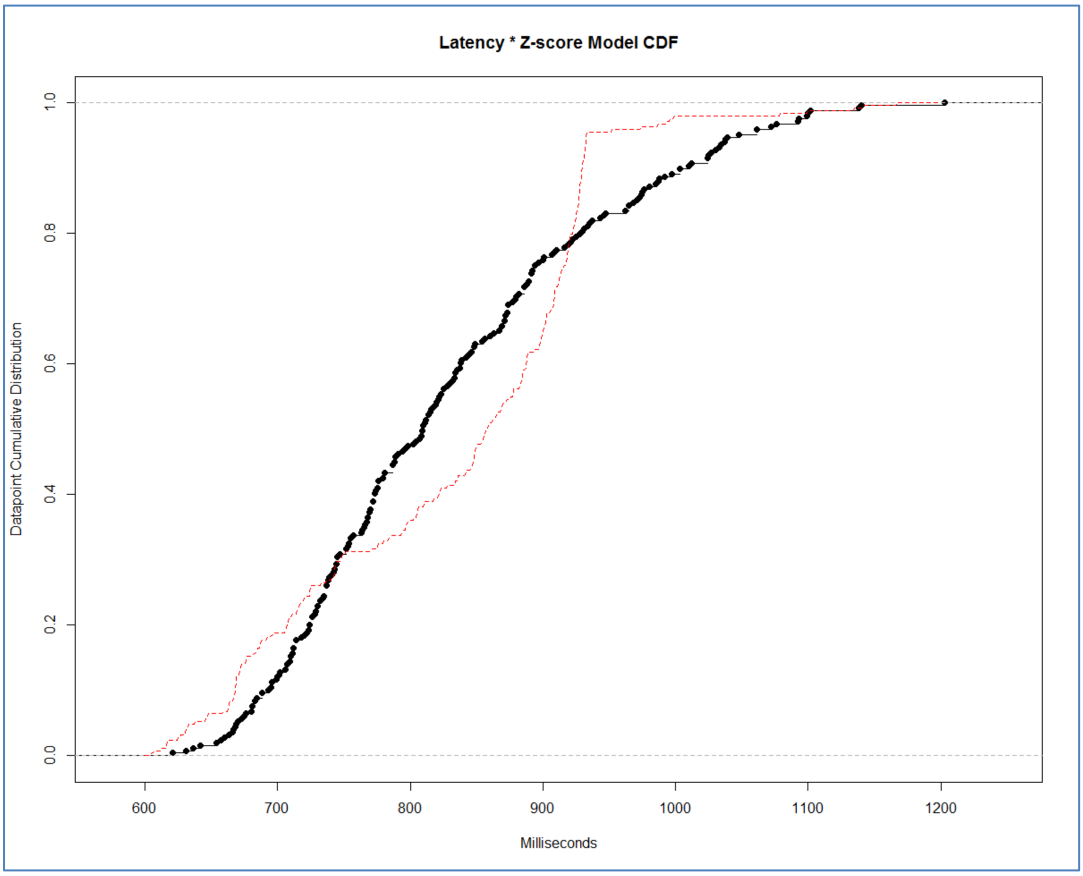
When depicted in a grain, the relationship between latency and Z-scaled latency can be more easily viewed (Figure 9, below). Estimating visually: 700 ms latency ≈ 780 ms Z-scaled; 800 ms latency ≈ 820 ms Z-scaled; 1000 ms latency ≈ 900 ms Z-scaled, and 1200 ms latency ≈ 950 ms Z-scaled.

Figure 9. Scatterplot of the Z-score model scaling, x-axis = latency, y-axis = Z-scaled Latency, linear regression line in red. The shaded area around the regression line represents the standard error.



When viewed as a pair of graphed CDF lines (Figure 10, next page), the Z-scaled latency observably matches the empirical CDF better than the L-scaled latency. Rather than two large divergent sections, it has three smaller divergent sections. The minimum Z-scaled latency begins at ~600 ms, ~20 ms before the minimum of latency, marking the first divergent section, running above latency from the minimum to ~30%, where the two lines intersect at ~750 ms. The second divergent section follows, running below the latency CDF, reintersecting with the latency CDF at the ~80% mark, denoting ~920 ms. The third and final divergent section follows, above the latency CDF, after which the Z-score model hits the maximum of its range at ~1050 ms, where ~98% of latency datapoints are measured, ~150 ms before the latency CDF hits its respective maximum.

Figure 10. Cumulative distribution function of latency (black dots), and Z-score model predicted latency (red dashed-line). Each dot/dash represents a single datapoint.



While this scaling method certainly is an improvement over the linear model, it does not utilize an ideally suited function either. Still, it is adequate for the current analyses. In the future, better scaling functions will need to be tested to generate closer correspondences between model and empirical data distributions.

7.2 **Visual Comparison**

N = 249	mean	SD	median	min	max	range	SE
Latency	828.36	118.45	810	621	1203	582	7.51
Cycle-time	25.61	2.04	26.25	21.13	30.77	9.63	0.13
Frequency	110.29	138.94	56	1	816	815	8.81
Levenshtein Distance	3.51	1.87	3	1	9	8	0.12
L-scaled Latency	828.36	48.62	843.79	721.67	951.42	229.74	3.08
Z-scaled Latency	828.05	111.77	857.28	604.65	1166.17	561.52	7.08

Table 7. Global condition.

N = 124	mean	SD	median	min	max	range	SE
Latency	809.05	111.59	777.50	621.00	1093.00	472.00	10.02
Cycle-time	25.67	2.01	26.30	21.13	30.37	9.23	0.18
Frequency	107.02	138.25	50.00	1.00	597.00	596.00	12.42
Levenshtein Distance	3.45	1.81	3.00	1.00	8.00	7.00	0.16
L-scaled Latency	829.85	47.95	844.86	721.67	941.92	220.24	4.31
Z-scaled Latency	831.33	110.06	859.92	604.65	1134.92	530.27	9.88

Table 8. Forward condition

N = 125	mean	SD	median	min	max	range	SE
Latency	847.51	122.33	828.00	631.00	1203.00	572.00	10.94
Cycle-time	25.54	2.07	26.25	21.18	30.77	9.59	0.19
Frequency	113.52	140.11	59.00	2.00	816.00	814.00	12.53
Levenshtein Distance	3.56	1.94	3.00	1.00	9.00	8.00	0.17
L-scaled Latency	826.87	49.42	843.79	722.66	951.42	228.76	4.42
Z-scaled Latency	824.80	113.80	857.28	606.35	1166.17	559.82	10.18

Table 9. Backward Condition.

When cycle-time & latency (descriptive data above, tables 7-9; further tables of descriptive statistics are found in the appendix, section 11.4) are represented as a histogram displaying the mean and standard deviation, the forward and backward conditions (Figure 11, next page) show a disparity: the model would appear to predict scarcely-faster translation in the backward condition than in the forward condition, whereas the latency data shows that the forward condition is faster than the backward condition by approximately 50 ms. The cycle-time bars should approximate the latency bars, but they do not, and opposite conclusions would be drawn from this comparison. But this is a comparison between cycle-time and latency. Perhaps the L-scaled and Z-scaled latency (Figure 12, next page), in category-view, show different results?

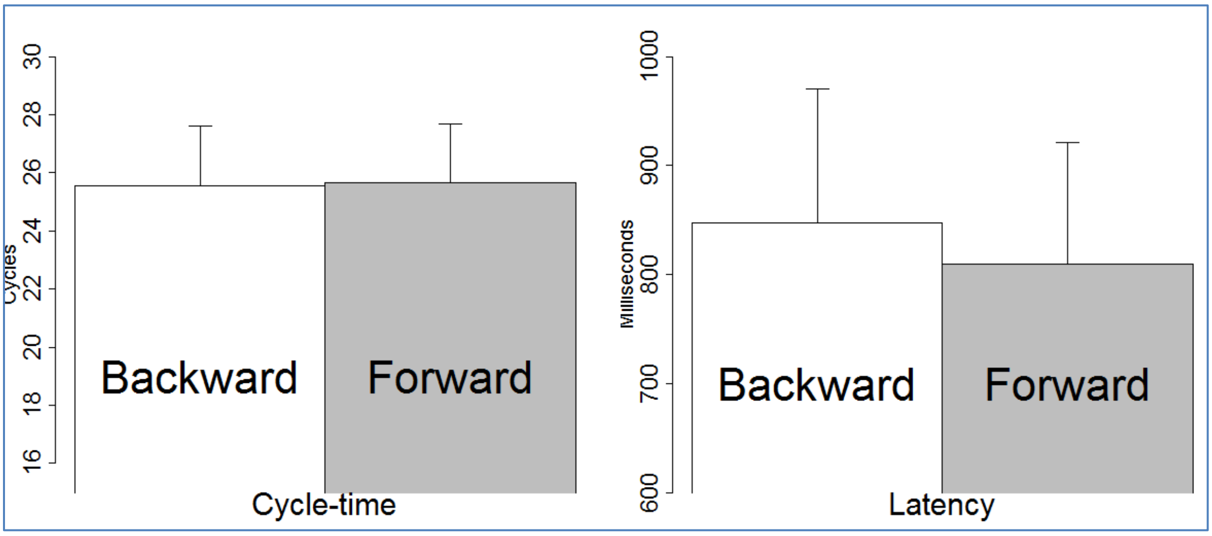


Figure 11. Histogram comparing mean cycle-time and latency for the backward and forward conditions. Barred-lines represent the standard deviation.

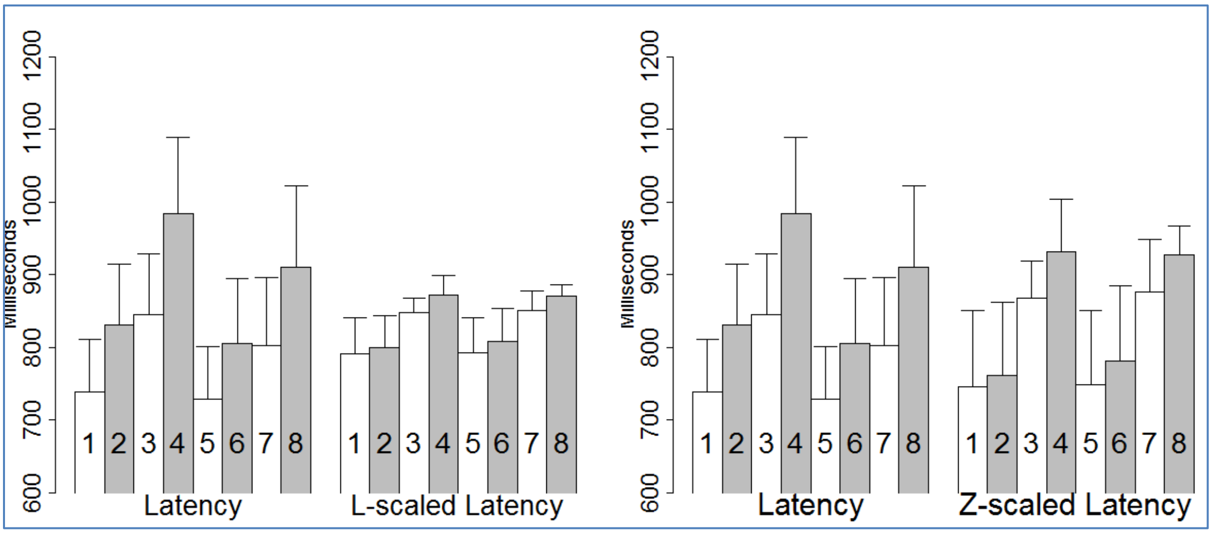


Figure 12. Histogram comparing mean latency, and the predicted-latencies, per category. Barred-lines represent standard deviation. Numerals: 1 = Backward high-frequency cognates, 2 = Backward low-frequency cognates, 3 = Backward high-frequency noncognates, 4 = Backward low-frequency noncognates, 5 = Forward high-frequency cognates, 6 = Forward low-frequency cognates, 7 = Forward high-frequency noncognates, 8 = Forward low-frequency noncognates.

The two graph-pairs show similar results. Figure 11 shows that cycle-time and latency give deviating results, with figure 12 corroborating this result for each category. Visually, neither predicted-latency measure seems to perform adequately with respect to the empirical data: L-scaled latency still suffers from a noticeably decreased range, which shrinks the means and the standard deviations, and produces a difference of ~100 ms between the cognate low-frequency and non-cognate high-frequency conditions which is not seen in latency. Z-scaled latency does not bear quite the same reduction in range, means, and standard deviations, but it calculates an even greater distance between the cognate low-frequency and non-cognate high-frequency conditions.

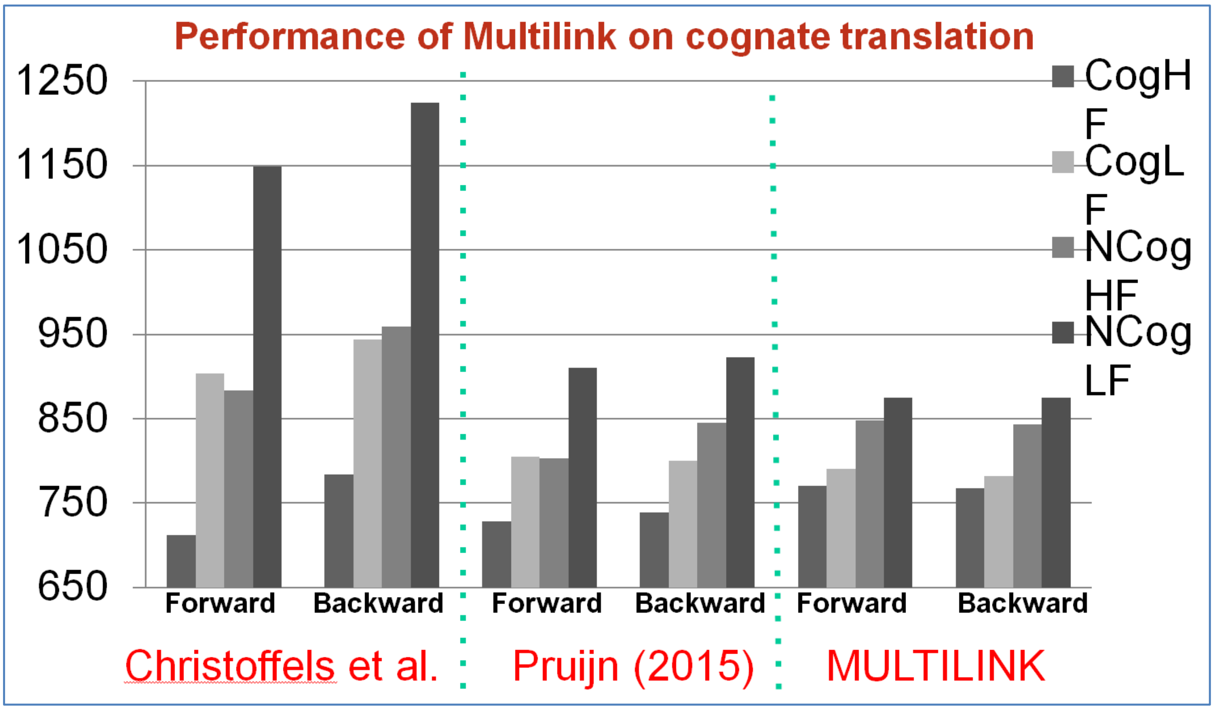


Figure 13. Performance Of Multilink On Cognate Translation. Data from Christoffels et al. (2006), Pruijn (2015), and Multilink (current thesis)

Based on the histogram comparison, Multilink does not perform well with respect to a number of translation aspects: there is no discernable translation direction effect, fundamental measures like mean and standard deviation are reduced, and the predicted-latencies do not all correspond to their empirical counterparts. However, these results also need to be situated in greater context, as shown in Figure 13 (above). Compared to the results of Christoffels et al. (2006), the difference between the results of Multilink and Pruijn (2015) are quite small. Both empirical studies tested the same variables (translation direction, cognate status, frequency, and proficiency). In fact, Pruijn (2015) is trying to replicate the results of Christoffels et al. (2006), even using many of the same stimulus, and both found a forward translation direction facilitation effect. Visually, this makes the results of Multilink appear quite satisfactory; a strong correlation between the empirical and model results can currently be expected.

But visual estimation is only the first step for determining the goodness-of-fit. For a more complicated view of the data, other analytical techniques must be used; their results are given in the following sections.

7.3 **Correlations**[18]

Empirical latency and predicted-latencies were tested for their correlation coefficients using the Spearman's Rank Correlation Coefficient test with the other manipulated numeric variables: cycle-time, frequency, and LD. The correlations can be computed for overlapping subsets of the stimulus. First, they can be computed across the entire 249-item stimulus set (global); second, they can be computed for each translation direction, cognate, and frequency condition groups separately, ~125 items per group; fourth, they can be computed across the 8 test conditions (cognate and non-cognate, High-frequency and Low-frequency, in both forward and backward translation directions), ~31 items per condition; and fifth, the 8 test conditions can be regressed to the mean, and correlations can be computed in this fashion. The following stimulus subsets have been correlated: global, forward & backward directions, 8 test conditions, and mean-regressed.

In all these cases, some correlations are more informative than others. In particular, we are interested in the correlations between latency, cycle time, and frequency (interactions with frequency are important, because this variable has been fundamental to the development of all available word recognition models). Based on the results of previous studies, some predictions can be made for these correlations: **latency by frequency**, and **predicted-latencies by frequency** should express significant negative correlations, while **latency by predicted-latencies** should express significant positive — nearly perfect, in fact — correlations. **Cycle-time by predicted-latencies** should be perfectly-correlated. **LD by frequency** should result in zero, or very small, correlations, but **LD** should result significant positive correlations for latency, predicted-latencies, and cycle-time (since cognates, with lower LD scores, are generally processed faster than non-cognates). The more specific conditions are predicted to show some other noteworthy correlations: for the non-cognates, frequency is predicted to show a higher degree of correlation than for the cognates; frequency is predicted to show a higher degree of correlation for high frequency groups than for low frequency groups. LD is predicted to show higher correlations with the cognate than the non-cognate conditions.

Correlations are presented as corregrams (Figures 14-18, following pages), visual matrix-representations of conditionally-related correlation coefficients, which have been colour-coded to indicate the strength of the correlation: blue for negative, and red for positive, with the colours becoming more saturated as ±1 is approached; colours at or near 0 are white. Rounded *p*-values are indicated underneath the test coefficient. Examination of the correlation test results, en masse, is found at the end of this section. For a complete

[18] $\alpha = 0.05$

tabled numeric index of the test coefficients and related output, refer to the appendix, section 11.5.

7.3.1 **Global Condition**

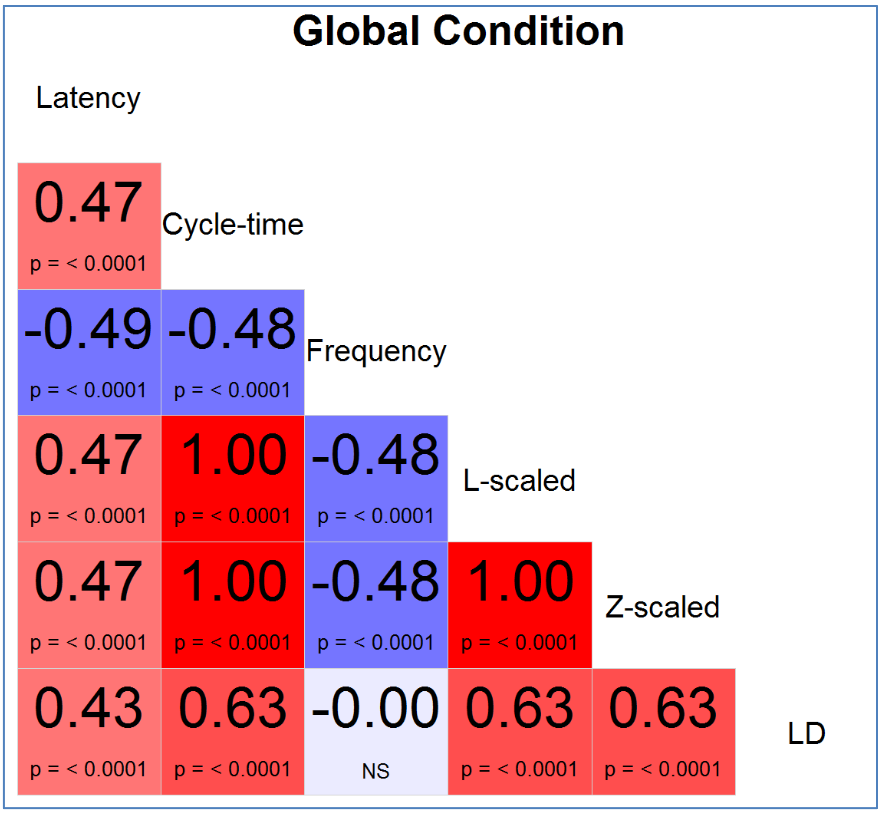


Figure 14. Corregram, global condition. "NS" = non-significant p-value.

When stimuli are globally correlated, the results are all highly significant (except LD * frequency), per the prior predictions.

7.3.2 **Forward & Backward Conditions**

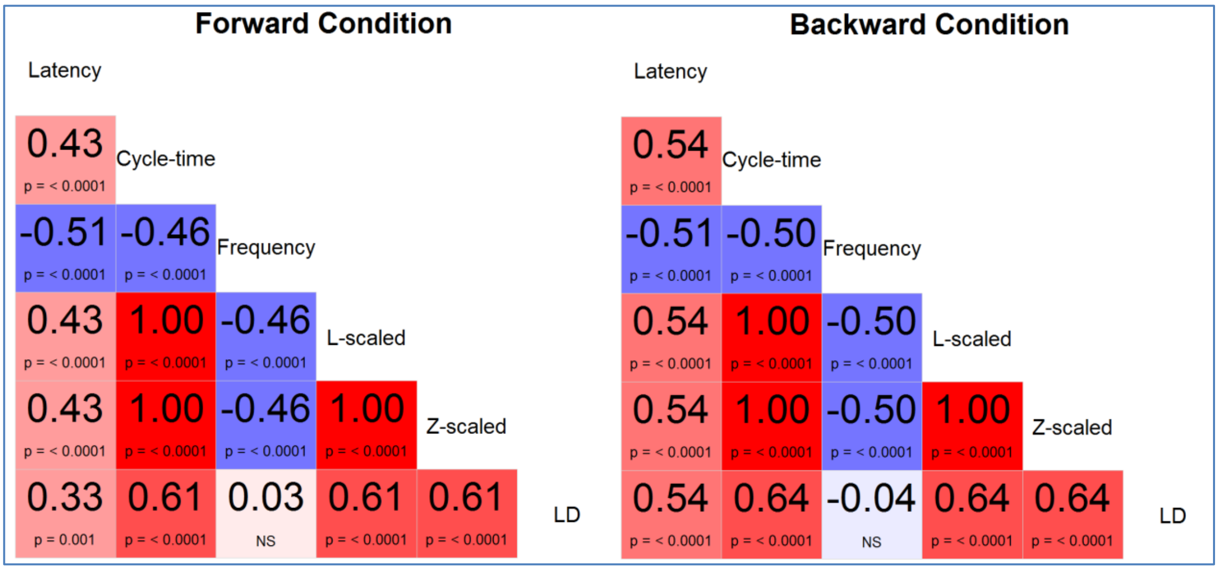


Figure 15. Corregram, forward & backward condition, combined for viewing convenience. "NS" = non-significant p-value.

Again, LD * Frequency is non-significant, but the other correlations meet expectations.

7.3.3 **Forward Conditions**

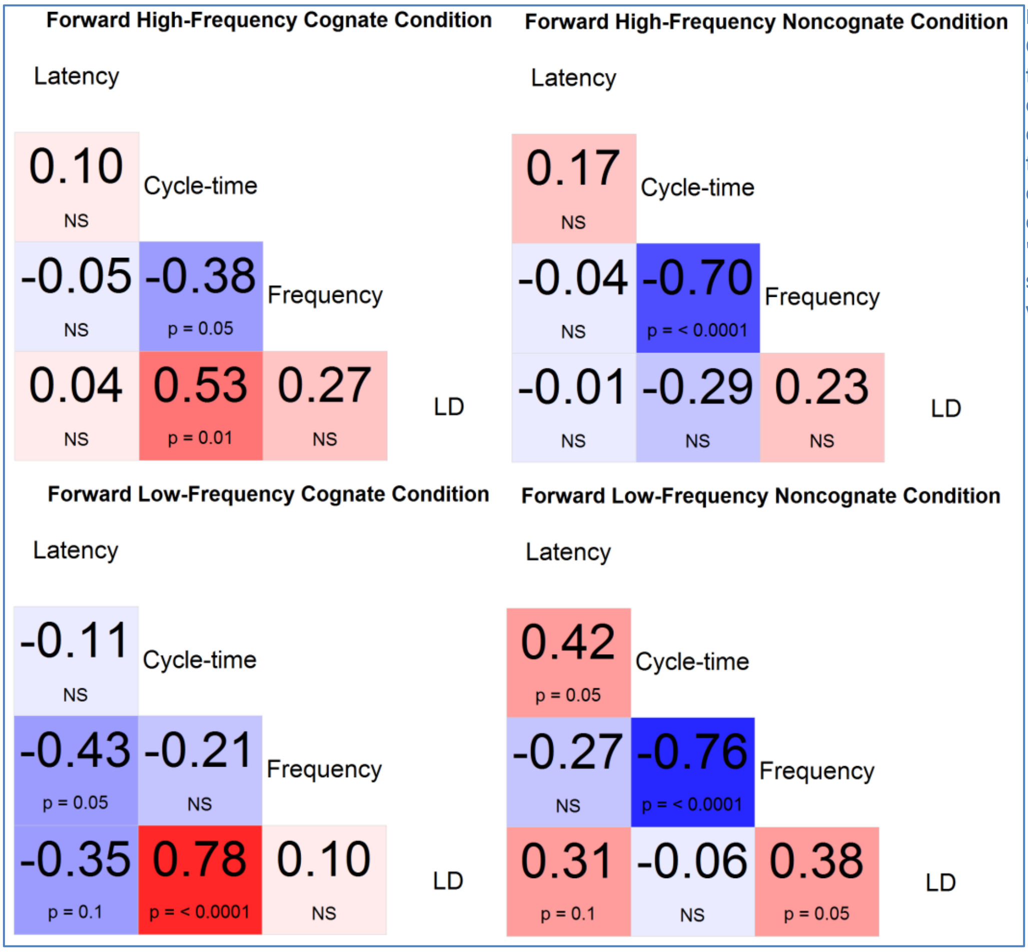


Figure 16. Corregram, forward conditions, combining frequency and cognate status categories. "NS" = non-significant p-value.

Due to the method of computation, cycle-time and the predicted-latencies have a perfect correlation, and exhibit the same correlations throughout the matrix. Therefore, they are deemed redundant, and removed from further analyses; cycle-time has been kept in their stead. The simplified corregrams facilitate the empirical (latency), model (cycle-time), and item (Frequency & LD) correlational analysis.

As observed in Figure 16 (above), many correlations are either totally non-significant, or above the alpha, Latency, and LD in particular.

7.3.4 **Backward Conditions**

Figure 17. Corregram, backward conditions, combining frequency and cognate status categories. "NS" = non-significant p-value.

Again, many *p-values* are observed to be non-significant, with Latency, and LD being almost totally above the alpha.

7.3.5 **Per-category Mean-regressed Condition**

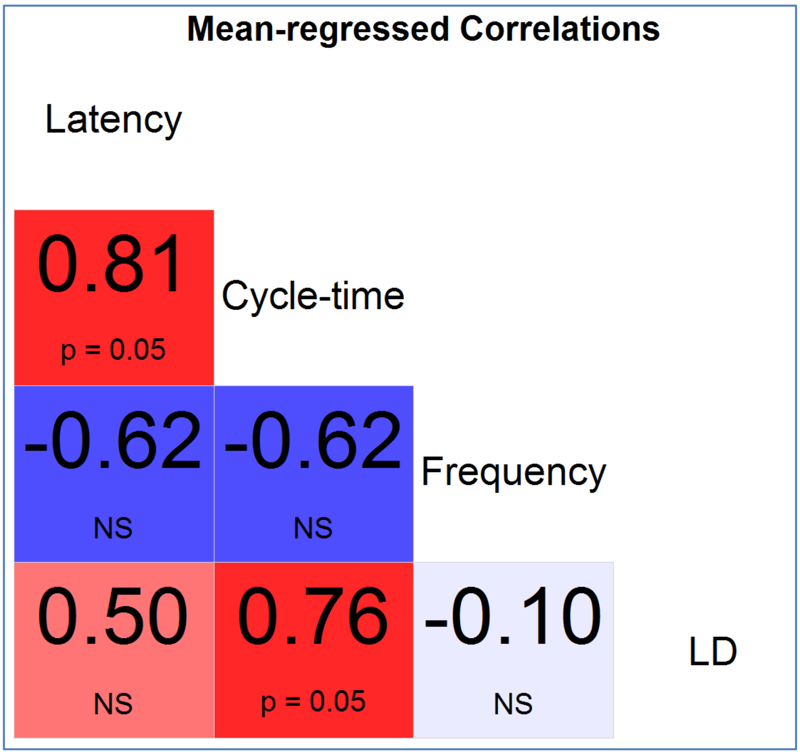


Figure 18. Corregram, mean-regressed condition. "NS" = non-significant p-value.

Mean-regressed correlations appear to suffer from the same non-significance issue. However, Latency * Cycle-time and Cycle-time * LD are noted to be strongly-correlated.

7.3.5 **Interpretations**

The correlational analyses exhibit interesting results. *Globally and directionally*, the correlations align per stated projections: LD shows significant positive correlations with the latency-group data, albeit an imbalance is observable between the LD coefficients of latency, predicted-latencies, and cycle-time. This imbalance could be considered indicative of the approach to candidate rating that Multilink takes, weighting LD to a greater degree than latency measurements actually exhibit. This imbalance is greater in the forward condition than in the backward condition. Frequency seems to be receiving similar correlative-coefficient scores for latency, cycle-time, and predicted-latencies. Unsurprisingly, the predicted-latencies, Z-scaled latency & L-scaled latency, are stated to be consistently perfectly correlated, both with each other, and with cycle-time, and show the same correlative-strength for every other tested variable combinations. Correlation testing does not demonstrate a performance difference between the two scaling models.

The *test conditions*, with smaller sample sizes, show a tendency for weaker correlations and non-significant *p*-values for latency, LD, and frequency. Comparing each set of correlations, two patterns become clear: first, frequency correlates with cycle-time\predicted-latencies to a greater degree in the non-cognate than the cognate conditions; second, LD only significantly correlates with the cognate conditions, and to a greater degree with the low-frequency cognates than the high-frequency cognates. *P-values* above alpha make other patterns more difficult to discern. Two results are intriguing: the weak significance and generally-low correlations seen with frequency, and the low correlations between latency & cycle-time, and, by extension also, latency & predicted-latencies in almost all of the experimental conditions. This can be attributed to the small sample sizes.

When the individual test conditions are *mean-regressed*, and correlations computed, the output results are slightly more positive. Latency by Cycle-time and Cycle-time by LD are seen to have a significant strong positive correlation. However, the other output correlations remain above alpha.

In total, most of the correlations were observed to either be weaker-than-expected, or non-significant. Conditional interactions are noted: backward translation direction correlations are stronger than forward translation direction correlations; latency and cycle-time are both strongly correlated with frequency for high-frequency conditions, and LD for cognate conditions, and the cycle-time particularly seems to correlate high-frequency or cognate with frequency or LD, respectively, whereas latency shows greater variability; the low-frequency noncognate conditions show that cycle-time correlates strongly with frequency, rather than LD, whereas latency shows medium correlations with both; conversely, the high-frequency cognate conditions show cycle-time correlates more with LD in the forward direction, but the backward condition shows near-equal cycle-time correlations for frequency & LD; latency, for the high-frequency cognate conditions, shows only a significant correlation with frequency in the backward condition. Theoretically, some implications can be drawn from these results: frequency possibly has greater empirical variability than previously expected; LD plays a larger role than previously considered, especially in the backward condition(s) and low-frequency noncognate conditions; Multilink, when cycle-time is mean-correlated, correlates with latency and LD quite well. Ultimately, however, it seems that correlational testing is not the best statistical method for discerning the relationship of these variables and conditions. Further testing was deemed necessary, requiring more advanced methods.

7.4 **Analysis Of Variance**[19]

ANOVA testing is employed to resolve a difference in effects. Which variables, when added to the statistical model, create a significant difference in the mean? The dependent variables being tested are: latency, cycle-time, and predicted-latencies. They are being

[19] $\alpha = 0.05$

tested against the independent variables: translation direction, frequency-category, cognate status, and all 2- and 3-way interactions.

The output F-values are beneficial to constructing a portrait of the comparative strength of tested variables, but these alone do not quantify effects in a standardized fashion. An effect size metric has been paired with these tests to help resolve the strength of each interaction. The Eta-squared (η^2) or Partial Eta-squared (η^2_p) effect size metric is commonly used to measure the strength of ANOVA outputs. The current study uses neither of these, selecting instead the bias-corrected partial omega-squared (ω^2_p) for effect size measurement, based on the conclusions of a recent study of effect size measurements (Okada, 2013).

Section 7.3.1 begins by comparing interactions with latency; section 7.3.2 compares interactions with cycle-time; section 7.3.3 compares interactions with L-scaled latency; and section 7.3.4 compares interactions with Z-scaled latency; section 7.3.5 discusses the results in total.

7.4.1 **Latency**

Interaction	DF	Sum Of Squares	Mean Square	F value	P-value (>F)	Partial Omega-squared
Direction	1	92094.25	92094.25	11.50	0.00081	0.04045
Cognate status	1	729728.00	729728.00	91.09	0.00000	0.26568
Frequency Category	1	668639.04	668639.04	83.46	0.00000	0.24879
Direction*Cognate status	1	24553.24	24553.24	3.06	0.08127	0.00822
Direction*Frequency Category	1	8479.03	8479.03	1.06	0.30461	0.00023
Cognate status*Frequency Category	1	24044.33	24044.33	3.00	0.08447	0.00797
Direction*Cognate status*Frequency Category	1	1070.71	1070.71	0.13	0.71499	0.00349
Residuals	241	1930670.60	8011.08			

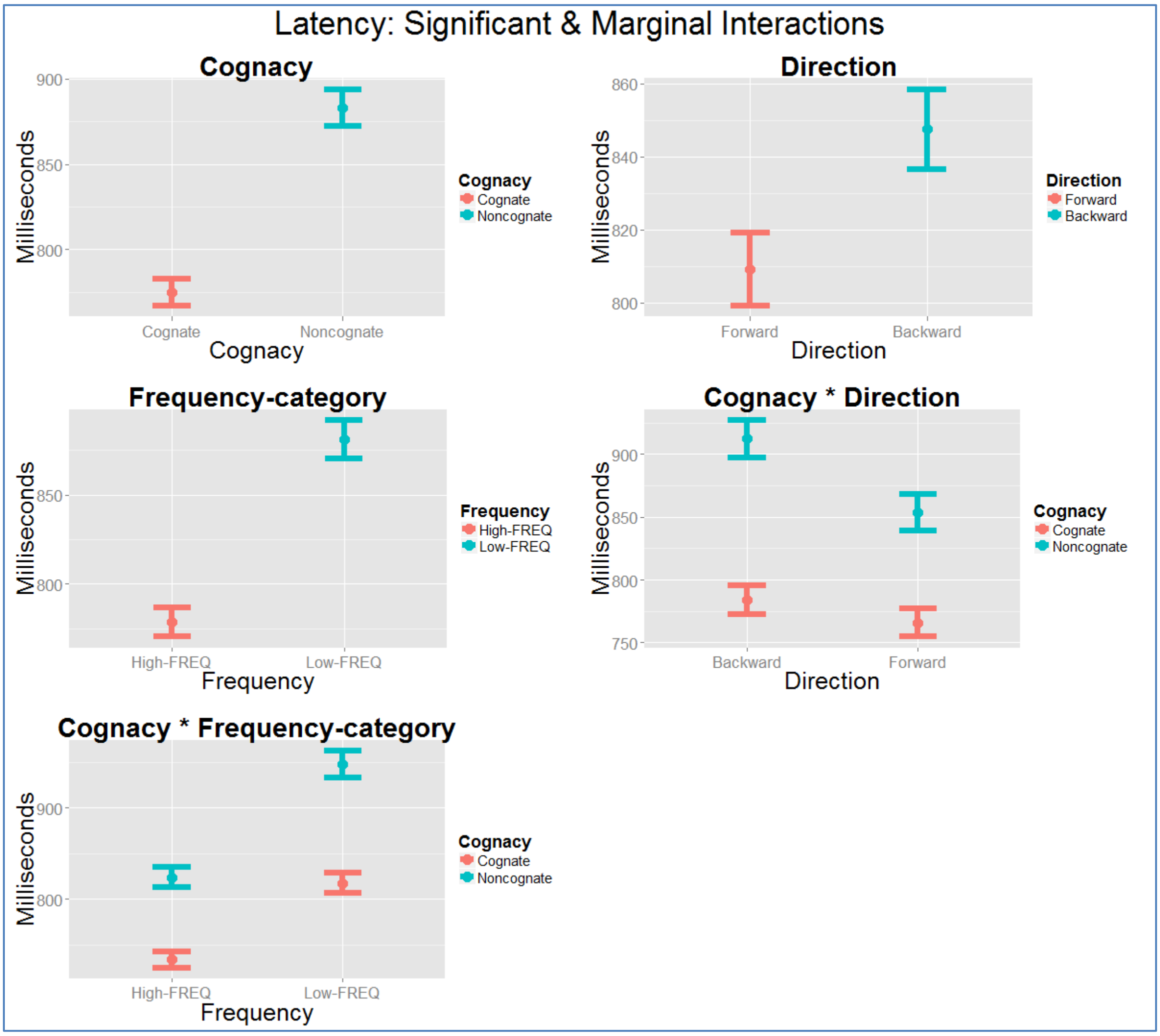
Table 10. ANOVA results for Latency * IV. Interactions above alpha are bolded.

Latency displays significant effects with **translation direction** (F (1, 241) = 11.50, p = 0.0008, ω^2_p = 0.04), **cognate status** (F (1, 241) = 91.10, p = < 0.0001, ω^2_p = 0.27), and **frequency-category** (F (1, 241) = 83.50, p = < 0.0001, ω^2_p = 0.25); marginal[20] interactional effects are observed with **translation direction * cognate status** (F (1, 241) = 3.11, p = 0.08, ω^2_p = 0.01), and **cognate status * frequency-category** (F (1, 241) = 3.00, p = 0.08, ω^2_p = 0.01). Insignificant effects include: translation direction * frequency-category (F (1, 241) = 1.06, p = 0.30, ω^2_p = 0.0002), and the 3-way interaction of translation direction *

[20] As stated in section 6.3.i, regressing latency measurements to the mean-per-token has reduced statistical power; these effects were found to be significant by Pruijn (2015: 23-25) in F1 ANOVA testing: "[. . .] a significant interaction between cognate status and translation direction was found (F1(1,36) = 233.46, p = .002, η^2p = .240; F2(1,245) = 3.37, p = .068, η^2p = .014) [. . .] an interaction between cognate status and frequency was found (F1(1,36) = 26.58, p = .000, η^2p = .43; F2(1,245) = 2.586, p = .109, η^2p = .010) [. . .]".

cognate status * frequency-category (F (1, 241) = 0.13, p = 0.71, ω^2_p = 0.004). Significant and marginal effects are graphed on the next page (Figure 19, below).

Figure 19. Significant and marginal interactions with latency; x-axis = respective independent variable categories, y-axis = mean latency, measured in milliseconds.



7.4.2 **Cycle-time**

Interaction	DF	Sum Of Squares	Mean Square	F value	P-value (>F)	Partial Omega-squared
Direction	1	0.97	0.97	0.41	0.52432	0.00239
Cognate status	1	416.51	416.51	174.10	0.00000	0.41009
Frequency Category	1	31.69	31.69	13.25	0.00033	0.04688
Direction*Cognate status	1	0.39	0.39	0.16	0.68778	0.00338
Direction*Frequency Category	1	0.12	0.12	0.05	0.82523	0.00383
Cognate status*Frequency Category	1	3.00	3.00	1.25	0.26387	0.00102
Direction*Cognate status*Frequency Category	1	0.99	0.99	0.41	0.52063	0.00236
Residuals	241	576.57	2.39			

Table 11. ANOVA results for Cycle-time * IV. Interactions above alpha are bolded.

Cycle-time displays significant effects with **cognate status** (*F* (1, 241) = 174.10, *p* = < 0.0001, ω^2_p = 0.41), and **frequency-category** (*F* (1, 241) = 13.25, *p* = 0.0003, ω^2_p = 0.05). Insignificant effects include: translation direction (*F* (1, 241) = 0.41, *p* = 0.52, ω^2_p = 0.002), translation direction * cognate status (*F* (1, 241) = 0.16, *p* = 0.69, ω^2_p = 0.003), translation direction * frequency-category (*F* (1, 241) = 0.05, *p* = 0.86, ω^2_p = 0.004), cognate status * frequency-category (*F* (1, 241) = 1.25, *p* = 0.26, ω^2_p = 0.001), and the 3-way interaction of translation direction * cognate status * frequency-category (*F* (1, 241) = 0.41, *p* = 0.52, ω^2_p = 0.002).

7.4.3 **L-scaled Latency**

Interaction	DF	Sum Of Squares	Mean Square	F value	P-value (>F)	Partial Omega-squared
Direction	1	553.46	553.46	0.41	0.52432	0.00239
Cognate status	1	236992.50	236992.50	174.10	0.00000	0.41009
Frequency Category	1	18033.37	18033.37	13.25	0.00033	0.04688
Direction*Cognate status	1	220.37	220.37	0.16	0.68778	0.00338
Direction*Frequency Category	1	66.53	66.53	0.05	0.82523	0.00383
Cognate status*Frequency Category	1	1707.31	1707.31	1.25	0.26387	0.00102
Direction*Cognate status*Frequency Category	1	563.37	563.37	0.41	0.52063	0.00236
Residuals	241	328067.41	1361.28			

Table 12. ANOVA results for L-scaled latency * IV. Interactions above alpha are bolded.

L-scaled latency, like cycle-time, displays significant effects for: **cognate status** (*F* (1, 241) = 174.10, *p* = < 0.0001, ω^2_p = 0.41), and **frequency-category** (*F* (1, 241) = 13.25, *p* = 0.0003, ω^2_p = 0.047). Insignificant effects include: translation direction (*F* (1, 241) = 0.41, *p* = 0.52, ω^2_p = 0.002), translation direction * cognate status (*F* (1, 241) = 0.16, *p* = 0.69, ω^2_p = 0.003), translation direction * frequency-category (*F* (1, 241) = 00.05, *p* = 0.83, ω^2_p = 0.003), cognate status * frequency-category (*F* (1, 241) = 1.25, *p* = 0.26, ω^2_p = 0.001), and the 3-way interaction of translation direction * cognate status * frequency-category (*F* (1, 241) = 0.41, *p* = 0.52, ω^2_p = 0.002).

7.4.4 **Z-scaled Latency**

Interaction	DF	Sum Of Squares	Mean Square	F value	P-value (>F)	Partial Omega-squared
Direction	1	2655.96	2655.96	0.37	0.54473	0.00255
Cognate status	1	1231477.09	1231477.09	170.58	0.00000	0.40513
Frequency Category	1	103091.19	103091.19	14.28	0.00020	0.05063
Direction*Cognate status	1	1046.64	1046.64	0.14	0.70372	0.00345
Direction*Frequency Category	1	108.25	108.25	0.01	0.90264	0.00397
Cognate status*Frequency Category	1	16197.92	16197.92	2.24	0.13547	0.00497
Direction*Cognate status*Frequency Category	1	3783.46	3783.46	0.52	0.46981	0.00192
Residuals	241	1739870.90	7219.38			

Table 13. ANOVA results for Z-scaled latency * IV. Interactions above alpha are bolded.

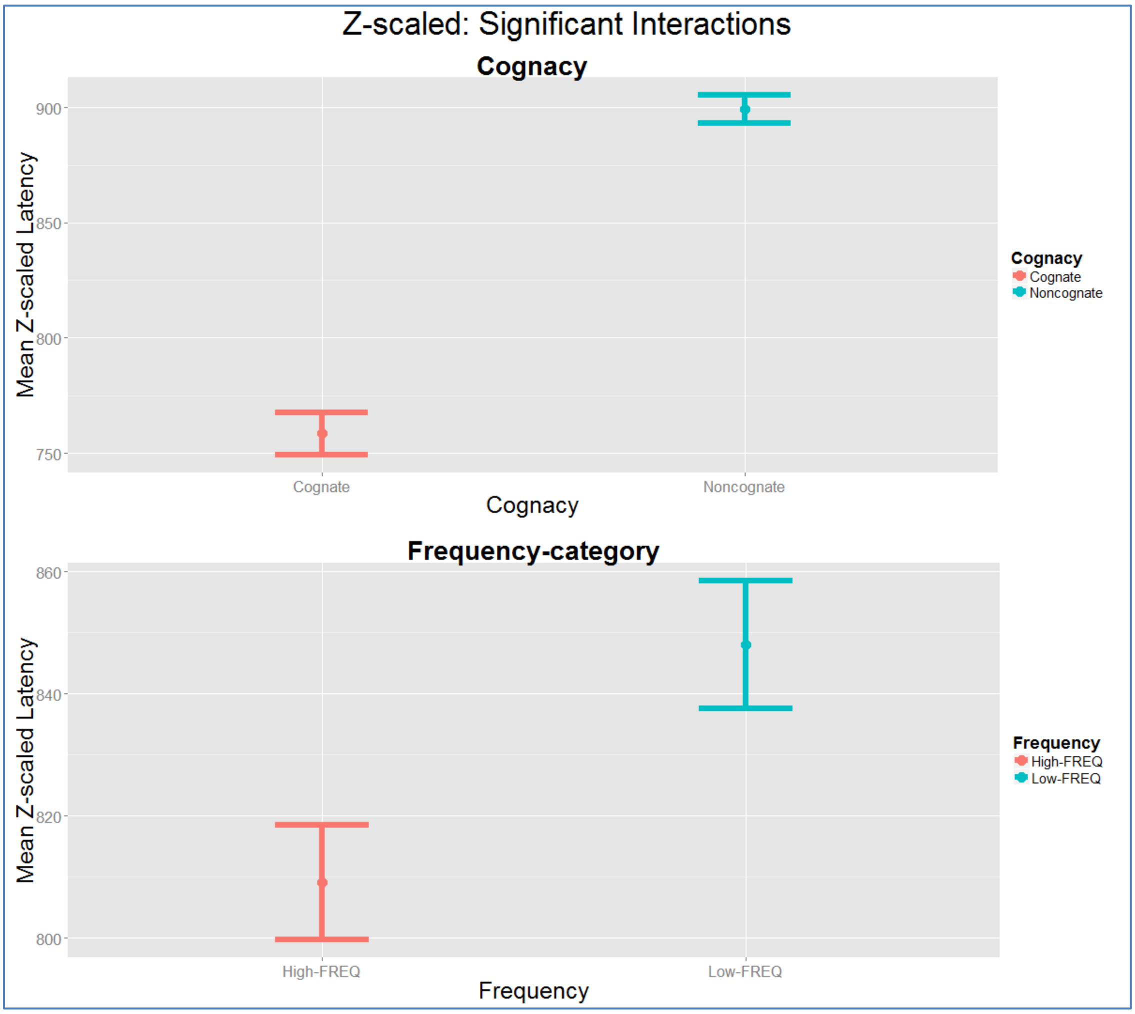


Figure 20. Significant interactions with Z-scaled latency; x-axis = respective independent variable categories, y-axis = mean Z-scaled latency.

Z-scaled latency displays significant effects for: **cognate status** (F (1, 241) = 170.58, p = < 0.0001, ω^2_p = 0.41), and **frequency-category** (F (1, 241) = 14.28, p = 0.0002,

ω^2_p = 0.05). Insignificant effects include: translation direction (F (1, 241) = 0.37, p = 0.54, ω^2_p = 0.003), translation direction * cognate status (F (1, 241) = 0.14, p = 0.70, ω^2_p = 0.003), translation direction * frequency-category (F (1, 241) = 00.01, p = 0.90, ω^2_p = 0.004), cognate status * frequency-category (F (1, 241) = 2.24, p = 0.14, ω^2_p = 0.005), and the 3-way interaction of translation direction * cognate status * frequency-category (F (1, 241) = 0.52, p = 0.47, ω^2_p = 0.002). Significant interactions are graphed in Figure 20 (previous page)

7.4.5 **Interpretations**

By examining the partial omega-squared effect sizes (including those without statistical significance, just for a full portrait), a depiction of the model's comparative performance can be established. Under this method of assessment, Multilink is found to be somewhat deficient in its replication of lexical effects.

Empirically, **direction** is rated as being a small, but determinable effect, but predicted-latencies underperform with respect to this variable by a factor of ~16x. Predicted-latencies overrate **cognate status**, on the other hand, by a factor 1.5x, although this variable does represent a very strong effect. **Frequency-category** is also empirically rated as a strong effect, but the model underestimates this by a factor of 5x. **Direction by cognate status**, and **cognate status by frequency-category**, while marginal under the current analysis, are known to be stronger when non-regressed (see footnote 14, page 60); still, neither are being well-replicated, and are currently underestimated by the predicted-latencies by a factor of 2.3x, and 1.6x, respectively. **Direction by frequency-category** is being overrated by the model by a factor of 17x. Finally, **direction by cognate status by frequency category** is underrated by the model by a factor of 1.7x. Generally, the Z-scaled latency has an effect size that is closer to the empirical effect size, except for the interactions of direction * frequency-category, and direction * cognate status * frequency-category, where L-scaled latency shows a closer-to-empirical effect size.

It is clear that the empirical data has a significant difference-of-means for 5 of the interactions (direction, cognate status, frequency category, direction * cognate status, and cognate status * frequency-category), and Multilink is not properly modelling these effects. Adjustments to the output mechanisms would seem to be necessary. But, significant noise factors could still be hidden within the data, affecting the test results. Predominantly, these could be bound in the controlled variables: phonetic onset, stimulus length, concreteness.

7.5 **Generalized Regression**[21]

Following from the ANOVA main effect tests, a regression model was constructed to check for interactions between main variables, plus their two-way interactions, and the controlled variables and their two-way interactions. Regression modelling is a method for estimating the relationships between a single dependent variable, and one or more independent variables, and their possible interactions. Latency and predicted-latencies were entered as the response variables, with direction, cognate status, frequency-category, phonetic onset, concreteness, and stimulus length functioning as the predictor variables. For the current data, the regression model will predict how latency or cycle-time is affected by the presence of independent and controlled variables, and their interactions. This will output a readily-interpretable estimation of the increase or decrease in milliseconds from the "grand mean", determine the amount of variance that is accounted for by the data collected, highlight latent effects stemming from controlled variables, and model non-parametric functions for the predictor variables. Each predictor variable plus its interactions is entered as a parameter into the regression model. Two-way interactions in particular were singled out in order to raise the true positive detection rate. Testing too many interactions would both reduce the true positive and increase the false negative detection rates by overfitting the model with more parameters than are actually necessary to fit the data. The output estimations for the empirical and model data should be comparable, and significant hidden empirical parameters — potentially not yet part of the model — should be observable.

A simple generalized additive model (GAM) was chosen over a linear model, linear mixed effects model, and a generalized additive mixed effects model, by selecting the model with the lowest amount of information gain\loss, measured by the Aikake Information Criteria (AIC). Regression-output and χ^2 Analysis Of Deviance tables are included with the results, however, due to the atypical number of tested interactions ($6^2 = 36$ possible model parameters), only statistically significant — including marginally-significant — results will be shown. Further tables are included in the appendix, section 11.XX. The intercept\regressed-mean for the model is defined as: backward, high-frequency, cognate, voiced consonant, non-concrete, short length. Significant and unexpected results will be visualized in accompanying graphs.

Section 7.4.1 details the latency results, section 7.4.2 discusses the L-scaled latency results, section 7.4.3 discusses the Z-scaled latency results, and section 7.4.4 combines and resolves these results into a full portrait of interaction.

[21] $\alpha = 0.05$

7.5.1 **Latency**

	Estimate	SE	Z Value	Pr(>|z|)
(Intercept)	700.7251	170.2812	4.115105	3.87E-05
Frequency-category: Low-frequency	176.3932	102.7258	1.717126	0.085956
Phonetic Onset: voiceless	244.4195	104.1739	2.346263	0.018963
Direction: Forward * Cognate status: Noncognate	-52.8048	24.79956	-2.12926	0.033232
Phonetic Onset: voiceless * Concreteness	-30.0559	16.2034	-1.85491	0.063608
N = 249	Model parameters = 28	Deviance explained = 0.521	R^2-adj = 0.46	UBRE\AIC = 111.36

Table 14. GAM significant and marginal interactions of Latency, and model information

Latency is predicted to have two statistically-significant group interactions outside of the regressed-mean: **forward noncognates** are estimated to be -53 ms below the regressed-mean (SE = 24.80, p = 0.03); and **voiceless consonantal onsets** are estimated to be 244 ms above the regressed-mean (SE =104.17, p = 0.02).

Marginally significant predicted effects: **low-frequency**, estimated to be 176 ms above the regressed-mean (SE = 102.73, p = 0.09); and **voiceless consonant onsets and high concreteness** scores are estimated to change mean latency by -30 ms (SE = 16.20, p = 0.06).

28 parameters have been selected to generate this model, accounting for 52.1% of the deviance in the dataset (R^2 = 0.46). Furthermore, according to the χ^2 test (table 22, below), deviance is significantly reduced by including phonetic onset (χ^2 (2) = 6.38, p = 0.04), and direction * cognate status (χ^2 (1) = 4.53, p = 0.03). Deviance is marginally reduced by including frequency-category (χ^2 (1) = 2.95, p = 0.09), and phonetic onset * stimulus length (χ^2 (2) = 4.86, p = 0.09). In particular, phonetic onset has the highest χ^2 score.

	DF	χ2 score	p-value
Frequency-category	1	2.948521	0.085956
Phonetic Onset	2	6.37513	0.041272
Direction*Cognate status	1	4.533768	0.033232
Phonetic Onset*Stimulus Length	2	4.863675	0.087875

Table 15. Analysis of Deviance latency test results. Larger coefficients attribute greater amounts of deviance to the respective predictor.

7.5.2 **L-scaled Latency**

	Estimate	SE	Z Value	Pr(>|z|)
(Intercept)	859.9264	69.10688	12.44343	1.52E-35
Cognate status: Noncognate*Stimulus Length	15.04643	4.763543	3.158664	0.00158494
N = 249	Model parameters = 28	Deviance explained = 0.536	R2-adj = 0.48	UBRE\AIC = 6.7668

Table 16. GAM significant and marginal interactions of L-scaled latency, and model information.

L-scaled latency is predicted to have only a single significant interaction above the regressed mean: **non-cognate**, when factored for **length**, is estimated to be 15 ms (SE= 4.76, *p* = 002) above the regressed-mean. There are no marginally-significant interactions between L-scaled and the stated predictor variables.

In addition, **cognate status * stimulus length** was found as the only significant predictor of deviance in Multilink's output (χ^2 (1) = 9.98, *p* = 0.002) (table 24, below).

28 parameters have been selected to generate this model, accounting for 53.6% of the deviance within the data set (R^2 = 0.48), when using L-scaled latency as the predictor variable.

	DF	χ2	p-value
Cognate status*Stimulus Length	1	9.977159	0.001585

Table 17. Analysis of Deviance L-scaled test results. Larger coefficients attribute greater amounts of deviance to the respective predictor.

7.5.3 **Z-scaled Latency**

	Estimate	SE	Z Value	Pr(>|z|)
(Intercept)	891.7014	159.9514	5.574829	2.48E-08
Cognate status: Noncognate*Stimulus Length	33.44751	11.02546	3.033661	0.002416058
N = 249	Model parameters = 28	Deviance explained = 0.531	R2-adj = 0.47	UBRE\AIC = 23.92

Table 18. GAM significant and marginal interactions of Z-scaled latency, and model information.

Z-scaled latency predicts the same interaction as L-scaled latency: **noncognate**, when **stimulus length** is factored, is estimated to lie 33 ms above the regressed-mean (SE=11.03, p = 0.002). There are no marginally significant interactions.

The χ^2 test (table 26, below) shows only one significant predictor of deviance: **cognate status * stimulus length** (χ^2 (1) = 9.20, p = 0.002)

28 parameters have been used to generate this model, which explains 53.1% of the total deviance in the data set (R^2 = 0.47).

	DF	χ2	p-value
Cognate status*Stimulus Length	1	9.203099	0.002416

Table 19. Analysis of Deviance Z-scaled latency test results. Larger coefficients attribute greater amounts of deviance to the respective predictor.

7.5.4 **Interpretations**

The results of the model-to-data comparison so far can be summarized as follows: latency has 4 significant and marginal predictors deviating from the regressed grand-mean: low-frequency, voiceless consonant onset, forward direction noncognate, and voiceless consonant onsets and concreteness. Conversely, the predicted-latencies only predict a single significant interaction, between cognate status and stimulus length. For both the empirical and model data, a 28 parameter model, consisting of the 1-way and 2-way interactions for the independent and controlled variables, is adequate to explain 51-53% of the deviance within the data (R^2 = 0.46 - 0.48). And whereas empirical data divides the explainable deviance between frequency-category, phonetic onset, direction * cognate status, and phonetic onset * stimulus length (with phonetic onset being the strongest predictor), the predicted-latencies assign similar amounts of explainable deviance to cognate status * stimulus length alone. Also note the rise in the intercept estimate: latency predicts approximately 700 ms, rising to 859 ms L-scaled, and 891 ms Z-scaled.

Focusing on latency, three results can be immediately eschewed: it is known that low-frequency words have higher latencies (see section 5.2.1.1, page 17), and the relationship between the forward direction and noncognate has already been established (see section 5.4.4, page 28); the significance of concreteness * voiceless consonant onset would seem to be a statistical confound, as there is no (current) proposition that correlates first-syllable onset phonemes with concreteness scores. The last interaction, between latency and voiceless onset (figure 21, next page) is, however, noteworthy.

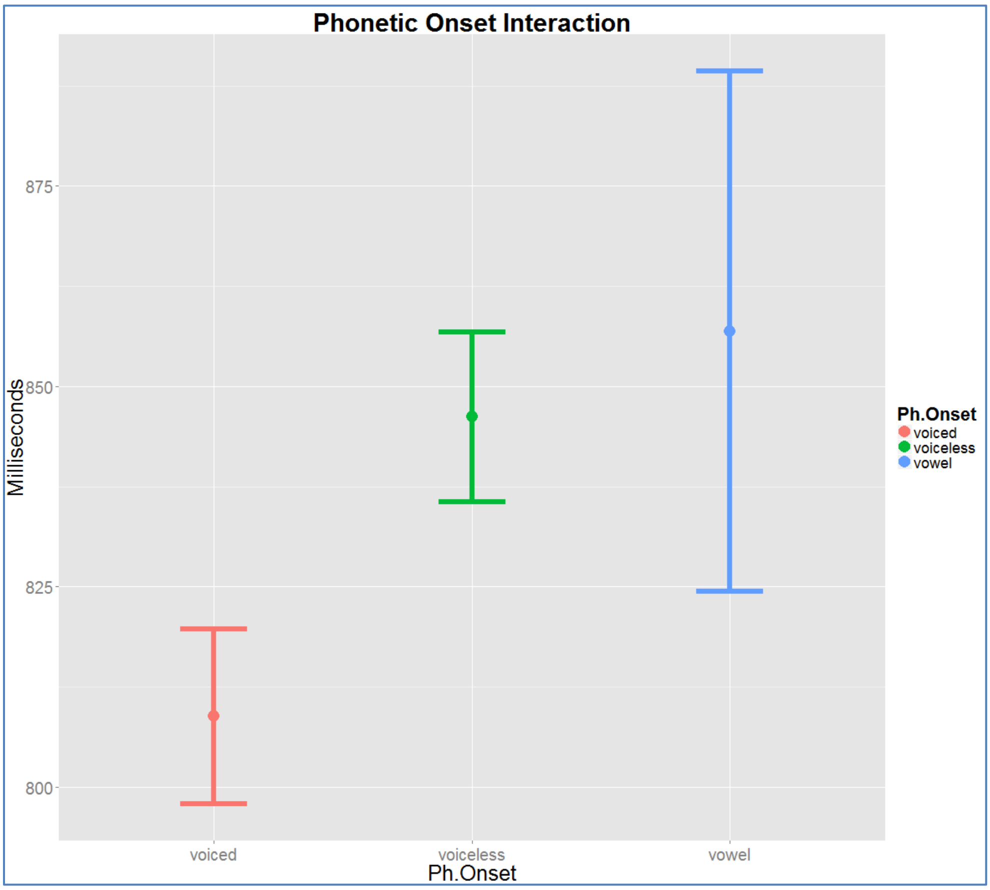
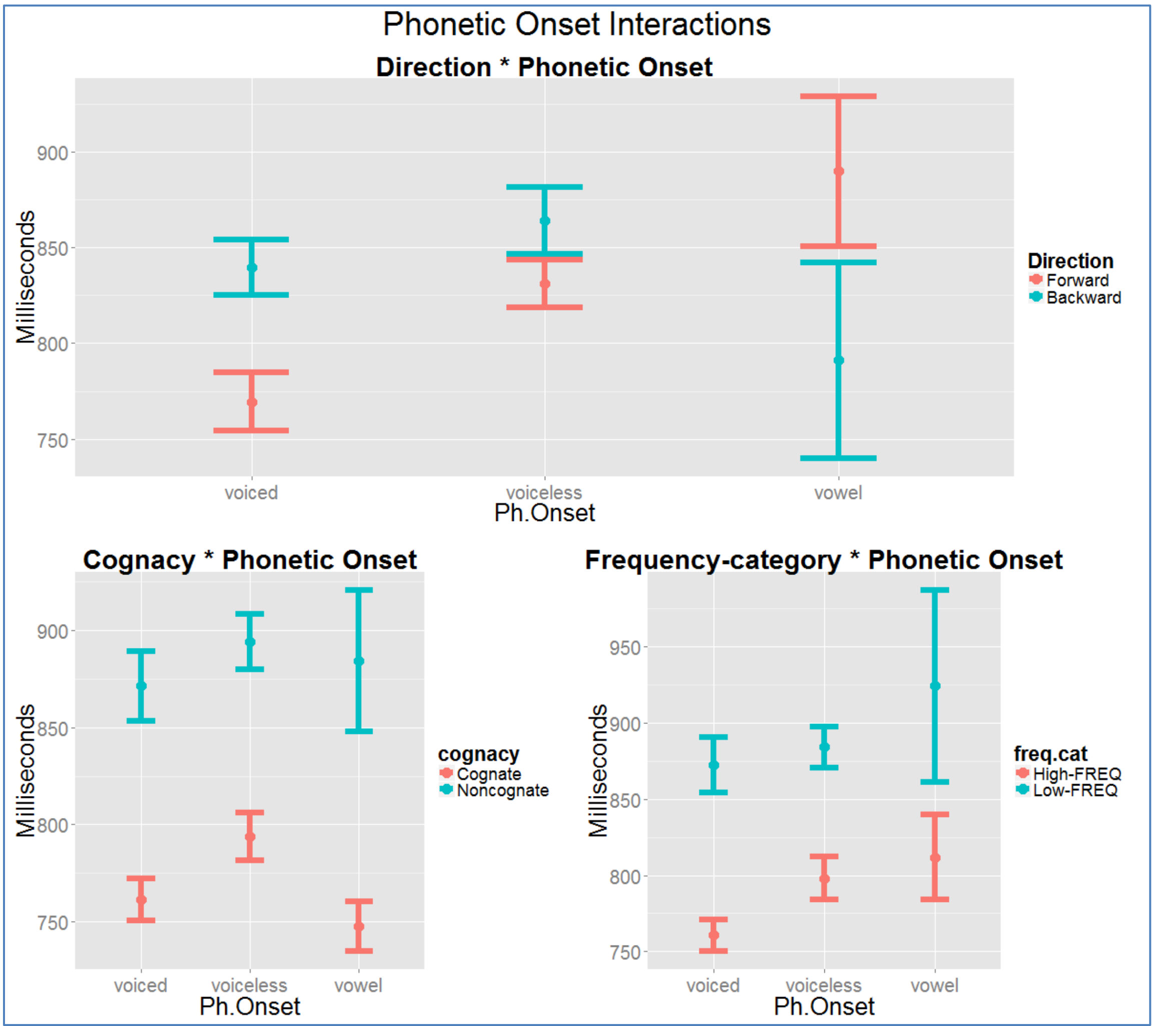


Figure 21. Interaction of phonetic onset with latency, x-axis = onset (voiced consonantal onsets in orange, voiceless consonantal oncsets in green, and vocalic onsets in blue), y-axis = mean latency, measured in milliseconds. Error bars represent standard error from the mean.

One likely explanation: the onset timing differences arise partly from differences in human articulation and phonation mechanisms, and also partly from the voice-key detection. This method of measuring participant production latency is not perfect, and, as noted in section 5.2.2.4 (pages 17-18), suffers from delays in recording sibilant, voiceless, and fricative consonants[22]. Nevertheless, if the voiceless onsets are discounted due to known articulation and phonetic factors, then the vocalic onsets are still unaccounted for; their standard error does not intersect at all with that of the voiced consonant onsets. Because the population of vocalic onsets in the stimulus is comparatively small, it is difficult to state definitively if there is a genuine effect present, and where it originates. This effect could vary systematically depending on the place and manner of articulation. Could direction be component to this, or perhaps cognate status and frequency (Figure 22, next page)?

[22] Theoretically, regression model output — like that found in this section — could be used to bias-correct stimuli with the particular onset properties. This method has advantages and disadvantages that will not be explored within this paper.

Figure 22. Potential interactions with phonetic onset, and the independent variables: direction, cognate status, and frequency-category. X-axis = voiced consonant, voiceless consonant, or vocalic onsets; y-axis = mean latency, measured in milliseconds.



Estimating visually, cognate status and frequency-category do not appear to interact with the phonetic onset. Direction, however, interacts in a peculiar fashion: in the forward direction, a linear rise is seen from the voiced onsets to the vocalic onsets; the backward direction has the opposite trend, showing vowels to have a lower latency than voiceless consonants (however, the standard errors here overlap significantly). Possible causes of this interaction will be explored in section 8.3. Multilink does not account for onsets, thus this data peculiarity is not seen in the model data.

One observation can be clearly stated: phonetic onset is a significant predictor of deviance, and combined articulatory and phonetic/phonological factors such as sibilance represent a very strong effect (stronger even than the cognate effect). But, it is not so strong that it obscures the main independent variable effects when balanced stimuli are utilized, and when onset type is not included in a statistical model.

The predicted-latencies show only one significant interaction: noncognates have significantly higher latencies when stimulus length is factored. This relationship is graphed below (Figure 23).

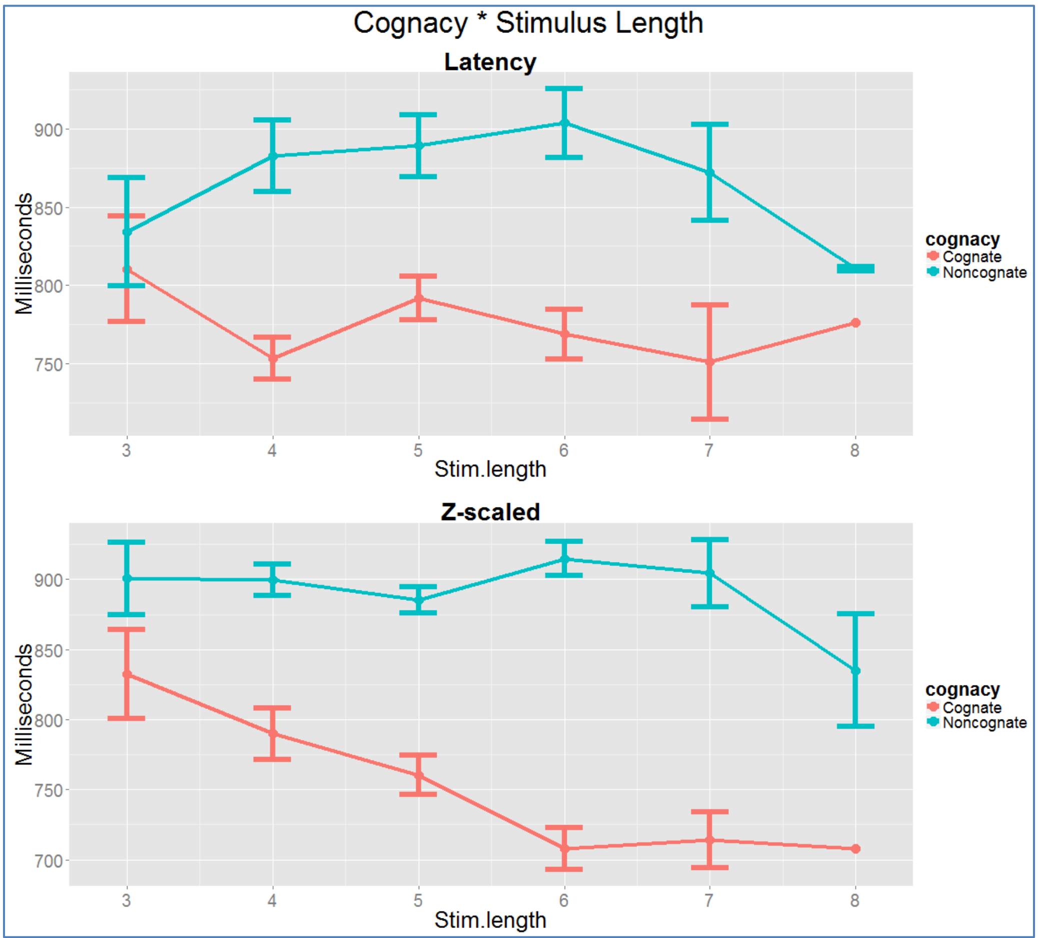


Figure 23. Latency & Z-scaled latency. X-axis = stimulus length, y-axis = mean latency, measured in milliseconds. Lines show the cognate groups. Bars represent standard error from the mean.

This interaction would appear to be partly misleading: it is no surprise that both the empirical and model data show a difference for cognates and non-cognates, both short and long. However, it is singularly conspicuous that the model and empirical graphs do not show any correspondence (if they did, then latency would also show that stimulus length * cognate status is a significant predictor of deviance in the $\chi 2$ test). Empirically, 3-letter cognates and non-cognates have ~40 ms difference in means (with significant SE overlap) in the 800-850 ms range; as cognates become longer, latency measurements vacillate between 750-800 ms. As noncognates become longer, latency rises, peaking at ~900 ms for 6-letter tokens, and then begins to descend. Multilink, Z-scaled, handles stimulus length somewhat differently: cognates have an approximately linear descent from ~830 ms (3-letter tokens) to ~720 ms (6-letter tokens), after which the measurements flatten. Noncognates 3-7 letters in length hover in the 880-920 ms range, before sharply descending at 8 letters to ~840 ms. There is no indication why the model differs in the treatment of cognate-groups, but it is clearly not a proper replication of empirical data. However, no tokens within the stimuli are above 8-9 letters in length. A second experiment using tokens 3-16 (or more) letters in length, compared to empirical measurements of the same tokens could be necessary to determine the adequacy of Multilink's cognate status * length output[23]. However, Multilink is currently also limited to an input maximum of 8-letters. This restriction would first need to be lifted.

Altogether, the regression model, combining the independent and controlled variables, revealed two important facets of the data that were previously hidden: phonetic onset is an important dimension that must be balanced due to phonetic and phonological elements that change the measured timing, but it has a potential interaction with translation direction (but likely not with frequency or cognate status); and Multilink processes cognates and non-cognates of increasing letter-length in a fashion that is not in accordance with current empirical measurements. In the empirical data, phonetic onset alone accounts for a large amount of deviance. Future modelling attempts will need to describe data for onset, and other potential phonetic factors. In order to rectify the current understanding, at least 2 new experiments to answer these questions might be necessary. The present data is not suited to generating a resolution, because the existing data is insufficiently described for phonetic and phonological components, which are already present and creating an onset measurement delay; and the input stimuli are length-restricted to 8 letters maximum.

7.6 **Divergence Testing**[24]

In order to make a final assessment concerning the "goodness-of-fit" that Multilink provides for the empirical data, a quantified comparison between the empirical and the model result distributions is necessary. This is accomplished via divergence (sometimes known as "statistical distance") tests, which measure the difference between two data groups. Different tests utilize varied base-measurements and manipulations, and many are based on theorems or proofs stemming from the field of information theory. "Total Variation Distance" and "Jensen-Shannon divergence" are two examples that are related to the "Kullback-Leibler divergence". Both scaling methods will be tested against the empirical data.

This section is divided into 3 subsections: 7.4.1 explains the use of the divergence tests, and particular idiosyncracies for the tests employed; section 7.4.2 shows the output for

[23] This could perhaps be accomplished using data from New et al. (2006), and tokens from the British Lexicon Project.
[24] $\alpha = 0.05$

each in a table, and 7.4.3 interprets these outputs in the global context of the data, and in the context of the other 2 tests.

7.6.1 **Test Details**

Dijkstra (1990: 163-165) provides specifications for a model-approximate χ^2 test (listed in the appendix, see section 11.2) that measures the distance of two data sets according to a χ^2 distribution, only requiring that the model data be transformed into analogous empirical data, already accomplished by the scaling methods described in section 7.1.

As part of a more general class of χ^2 tests, it is unbounded, built upon sample variance testing, and requires independent, Gaussian-distributed observations that can be linked to a χ^2 distribution function. The *null hypothesis* assumes that the model distribution can approximate the empirical distribution; however, this type of test requires large sample sizes, and does not work well with small samples (sample size must at least satisfy the central limit theorem). Essentially, as with a Pearson χ^2 test, it determines whether there is a significant difference between the sample variance of observations in the empirical distribution, and the model distribution. Given this knowledge, a *high p-value* is expected, showing that the null hypothesis — that the two distributions match to a significant degree — is true. The degrees of freedom, which determines statistical significance, is generated by subtracting the number of model parameters (the number of dimensions that fit the data) from the number of observations (ie sample size). *P*-values have been estimated using a critical value distribution table.

Because the model-approximate test is relatively unknown and seems infrequently-applied, the results will be supported and compared to 2 commonly-accepted and established divergence tests: the Kolmogorov-Smirnov statistic, and the Kullback-Leibler Divergence test, both of which can assess the model-fit to the empirical data by "matching" the distributions formed by both samples. Furthermore, these additional tests were chosen due to issues with data-pooling in the model-approximate test, which reduce the statistical power and can potentially lead to excess rejections.

The Kullback-Leibler Divergence (D_{KL}) (Kullback & Leibler, 1951) test is not a true metric, but rather a "premetric" based on the space of probability distribution functions for the input data. It is asymmetric [$D_{KL}(P||Q) \neq D_{KL}(Q||P)$][25], therefore the test must be run twice, and input data must be averaged from both sides (P||Q and Q||P give different results) in order to obtain the "symmetric" result. The output statistic is unbounded, having a theoretical

[25] For the non-mathematically-inclined, this is read as: "The KL-divergence of P on the condition of Q does not equal the KL-divergence of Q on the condition of P."

limit of ∞, always non-negative, and represents a measurement of the information gain when making the first input distribution function approximate the second. Numbers closer to 0 show less gain between distributions. This output is typically in Shannons/bits, but can be output as "nats" (natural-log base), or "hartleys" (decimal base). It is also known as "relative entropy", being related to other information measurements such as self-information and Shannon entropy. The *null hypothesis* assumes that both distributions share approximately the same probability distribution space. Outputs above alpha show that the distributions are significantly different from each other. It is often used in the field of machine learning and artificial intelligence as a criterion for feature extraction and model selection.

The Kolmogorov-Smirnov test[26] (*KS*) (Kolmogorov, 1933; Smirnov, 1948) is a nonparametric test that quantifies a sampled distance between two cumulative distribution functions (as seen in figures 10 & 12). Originally designed for comparison with reference distributions (gaussian, gamma, or poisson distributions, for example), a one-sample KS-test, a two-sample variant for goodness-of-fit testing has been in use for quite some time, finding a place in the analytic techniques of astronomy (Peacock, 1983), and physics (Lopes, Reid, & Hobson, 2007). The *null hypothesis* for the two-sample KS-test assumes that both samples have equal cumulative distributions, and therefore are derived from the same population or distribution. It is sensitive to divergences of both location and shape, thanks to an inherently high sample-rate. This high sample-rate is also its weakness, as it works best with large datasets, otherwise generating possible false positives. The null hypothesis is rejected if the output statistic — a measure of the absolute maximum distance between the two samples — is larger than the chosen α; the closer the statistic is to 0, the more probable that both inputs have the same distribution. The output is bounded, $KS \leq 1$, and symmetric.

Both supporting tests are accepted as suitable statistics for determining "reality vs model". Accordingly, true experimental measurements within the input are a theoretical requirement of both tests. Opposite this, the requirement for experimental data is directly built-into the model-approximate χ2 test.

[26] Some potential issues with the KS test are outlined here: https://asaip.psu.edu/Articles/beware-the-kolmogorov-smirnov-test

7.6.2 **Output Statistics**

Model \ Test	Model-approximate χ2 Test	Kullback-Leibler Divergence (symmetric)	Kolmogorov-Smirnov Statistic
Linear Model	3.56 (0.99)	0.012	0.28 (< 0.0001)
Z-score Model	2.73 (0.99)	0.016	0.18 (0.0008)
	DF = 247	Parameters = 2	N = 249

Table 20. Divergence test values. p-values in parentheses. KL-divergence does not output a p-value.

For the linear model, χ^2-approximate (247) = 3.56 (p = 0.99), requiring us to accept the linear model as being adequately-fitted towards the empirical data. *KS* = 0.28 (p = < 0.0001) bits, rejecting the null hypothesis that both distributions come from the same population; the observed information gain is very small, however, showing that although these distributions are not derived from the same population, they are close approximations. D_{KL} = 0.012 bits, showing a small amount of information gain.

The Z-score model shows slightly better results: χ^2-approximate (247) = 2.73 (p = 0.99), requiring us to accept the Z-score model as being adequately-fitted towards the empirical data. *KS* = 0.18 (p = 0.0008) bits, rejecting the null hypothesis that both distributions come from the same population; but this is a very small amount of information gain, showing that the two distributions are nearly approximate. D_{KL} = 0.016 bits, a slight gain in information.

For reference, the divergence of cycle-time & L-scaled latency is D_{KL} = 0.0003 bits. Similarly, the divergence of cycle-time & Z-scaled latency is D_{KL} = 0.002 bits

7.6.3 **Interpretations**

The previous tests have measured the empirical-to-model statistical divergence, based on information-theoretic probabilities, approximating CDFs, or sample and model variance. Results are positive, showing a low degree of empirical\model divergence, indicative of why Multilink performs so well: though its replication is not perfect, it does manage to fit the empirical curve in an apparently satisfactory way. There is little information gain between the empirical CDF, and the model CDF. By converting the χ^2-approximate scores into Shannons\bits, like the others, the 3 values can be directly compared: L-scaled diverges by 0.024 bits from the empirical data, and Z-scaled diverges by 0.032 bits from the empirical data. As a result, the model has a final divergence rate of maximum 0.032 bits, and minimum 0.012 (range: 0.020), less than a single bit, even at the worst. Curiously, the linear model is predicted as having fewer gained bits than the Z-score model by the KL-divergence test and the χ^2-Shannon score (although the critical-value is rated lower for the Z-scaled latency). The reason for this might be due to the fact that the linear model is constructed around the central tendency of latency and cycle-time, and predicting the L-scaled latency data based on the distance from this central tendency. The Z-score model, conversely, is constructed upon a Z-value and logarithmic transformation of individual data-points, which are then used to predict the Z-scaled latency. If this is the case, then it would appear that the KL-divergence test and χ^2-Shannon are picking up on this difference.

It is worth noting that this model has been applied to Dutch and English words, two very similar West-Germanic languages within the Indo-European family. Multilink has been indirectly optimized for use with these two languages, the language family, and also with the latin script, in the same way that other signal transmission and processing systems can be optimized for use with a specific alphabet[27]. It is an inherent, if unconscious, part of the design. Inevitably, it will need to have a language of distant typological classification (which should show a higher cycle-time for the translation production task, because increasingly distant languages become harder to adapt into the mental lexicon (Schepens, 2015)) added into the lexicon to test this assumption. Nonetheless, Multilink has an accuracy rate of just over 98% despite any current shortcomings. The present architecture allows it to sufficiently replicate the word-translation process.

7.7 **Test Outcomes**

At last, the results of this study are generally positive. The scaling methods (section 7.1) used — linear model scaling, and Z-score model scaling — were employed to transform cycle-time data into approximate-millisecond measurements, and then later tested against each other in subsequent sections to find which fits the empirical data better. Z-score model scaling was typically found to be superior throughout testing, as the linear model suffers from a greatly reduced approximate-millisecond range, although the performance differences did not show a major preference for one over the other.

A visual comparison of the model and empirical results (section 7.2) using histograms shows a forward translation direction effect (forward translation is ~40 ms faster

[27] Morse code is one example of this (Çiçek & Yilmaz, 2013)

than backward translation, the opposite prediction of the RHM) in the latency data that is not replicated by the model (Figure 13). Per-category, combining the binary independent variable groups of cognate status, frequency-category, and direction, (Figure 14) shows that Multilink fares better in some categories than others; in particular, it produces a gap between the low-frequency cognates and the high-frequency non-cognates that is not seen in the empirical data. This was seen as a first indication that Multilink is not properly simulating, possibly overrating, cognate status & frequency in the output.

Correlational analysis (section 7.3) using Spearman's Rank Correlation was rather unsuccessful at divulging particularly noteworthy information. 12 conditions were tested, starting from N = 249 (all stimuli), to N = 128 (direction categories), to N = 32 (the independent variable categories), and N = 8 (mean-regressed), with latency, cycle-time, frequency, L-scaled, Z-scaled, and LD being tested for correlations (L-scaled & Z-scaled, the predicted-latencies, were removed for the category correlations because they displayed the exact same correlation coefficients as cycle-time, making them redundant). While predictions were successful for the conditions with larger sample sizes (the global, forward, and backward conditions), the specific categories suffered from problems of statistical significance. Surprisingly, many correlations, even those with statistical significance, were found to result in weak coefficients, and little interpretable information was uncovered. The mean-regressed correlations did at least demonstrate that, when averaged, Multilink maintains a strong positive correlation with latency, and LD.

Where correlational analysis failed to reveal much, ANOVA testing (section 7.4) was conducted to determine statistically significant interactions between the input response variables (latency, cycle-time, and the predicted-latencies), and the predictor variables (translation-direction, cognate status, and frequency-category), and all 2-way and 3-way interactions. Latency was found to have 3 significant predictors and 2 marginally-significant predictors ("marginal" at $\alpha = 0.1$ rather than the afore set $\alpha = 0.05$, the general significance threshold). Predicted-latencies were found to have only 2 significant predictors. Effect size metrics were then used to determine that Multilink is overrating the strength of the cognate effect compared to empirical data, and is underrating the effects of direction and frequency. The need for parameter-refinement was established based on this test.

In order to tease out any hidden effects deriving from the controlled variables (phonetic onset, concreteness, and stimulus length), and determine the total amount of explainable deviance within the collected data, a generalized additive regression model (GAM) (section 7.5) was selected from other comparable models, including a generalized linear model, decided based on a lower Aikake Information Criterion score. The aforementioned controlled variables were tested, along with the main independent variables, plus all 2-way interactions (but not 3-way interactions) were tested as the predictor variables. Latency, cycle-time, and predicted-latencies were used as the response variables. Collected data accounts for approximately 52% of explainable deviance. Empirically, latency assigns the largest amount of deviance to phonetic onset, an issue arising from unaccounted phonetic and phonological factors that affects the word-naming measurements. It is not so strong that it obscures other effects, but future studies will need to consider these additional factors. A curiosity was noted for direction * phonetic onset interaction, currently unresolved. Multilink's only significant predictor of deviance was an interaction of cognate status * stimulus length, possibly arising from the output ranking scheme, which length-normalizes LD.

Lastly, divergence testing (section 7.6) was done, using 3 different divergence statistics: a model-approximate $\chi2$ test, the (symmetrized) Kullback-Leibler Divergence test, and the Kolmogorov-Smirnov statistic. These tests showed somewhat mixed results between the two scaling models, but did indicate that the model performs overall very well with respect to the empirical data.

This indicates why Multilink is able to have such a high accuracy, despite any deficiencies: it replicates the empirical data curve to a satisfactory degree. Highly-salient lexical dimensions and their facilitatory or inhibitory effects — which subsequently precipitate greater alternations in the results distribution — are represented by the model. So, it can be stated that, although Multilink is not a perfect simulator of the human cognitive process of lexical access, for a first-generation computational model of bilingual recognition & production, it has stellar performance. It is in need of some minor corrections, particularly in light of the findings of Pruijn (2015), but despite this, it sets a high bar for competing bilingual models in the future.

8 Discussion

The following sections discuss the results of this thesis in light of the literature already reviewed, integrating them into our current understanding of bilingual lexical access and processing, and advancing possible loci for future research. Section 8.1 discusses the overall performance of Multilink in its current state, compared to the expectations of the RHM and the BIA+; section 8.2 reexamines the facilitation effects, and how the most important dimensions can be adjusted within the model; section 8.3 discusses the latent interactions.

8.1. **Performance Of Multilink**

As found throughout the results section (section 7) and the previous studies (Dijkstra & Rekké 2010; Dijkstra et al., in prep.), Multilink version 1.02 is thus-far an imperfect, but promising, psychometric task simulator.

With respect to the simulated experimental data, word-naming latency has 6 detectable effects of varying strengths, listed in order of descending strength: phonetic onset effect, cognate facilitation effect, frequency facilitation effect, forward translation-direction facilitation effect, forward translation-direction facilitation effect & cognate facilitation effect interaction, and the cognate facilitation effect & frequency facilitation effect interaction. Multilink replicates only two of these adequately, and an additional interaction: the cognate facilitation effect, the frequency facilitation effect, and a cognate & word-length effect interaction. Building upon previous models, particularly the RHM and BIA+, Multilink has inherited their strengths and weaknesses. The nested BIA+ architecture focused on cognate & noncognate detection, forming the core of the model. This design makes for fast and sufficient detection of bilingual words, an empirical reflection of cognate effect activity, but leaves the model prone to errors when tokens lack semantic representation, which seems to be used as a final "pinpoint" measure for selecting the translation equivalent when candidates are being ranked equally by their LD. As stated in section 7.4.5, following ANOVA testing, Multilink overrates the power of the cognate effect compared to empirical data. The frequency effect is not adequately replicated either, being underrated as an effect by the model. Empirically, the frequency effect is almost as influential as the cognate effect in altering production latency; highly-frequent noncognate tokens have nearly the same latency as low-frequency cognate tokens. The lack of a translation direction effect can be pardoned: empirical evidence has been disputative and inconclusive until the results of Pruijn (2015). If anything, a potential backward translation direction effect is observed[28], but the effect is not statistically significant in the current simulations. This would be in line with

[28] Noted in Dijkstra, et al. (forthcoming: 45) for "low-proficiency" modulations (ie, word form frequency OPM divided by 4 to reflect lower subjective frequency) within Multilink.

the estimates of RHM, which has hypothesized backward translation direction facilitation due to reliance on L1-associative links from the L2, rather than direct linking to the concept/semantic representation both connections interdependent upon L2 proficiency. While the RHM has remained the dominant model in bilingual lexical access (section 5), recent empirical studies (Kroll et al., 2002; Christoffels et al., 2006; Pruijn, 2015) have shown differing outcomes, so it is no surprise that Multilink precludes this effect without first altering the input frequency. The remaining interactional effect, cognate status & frequency (translation-direction by frequency can be discounted due to the lack of directional facilitation), should be significant in the output of the model, but is not rated as being particularly active. The lack of interaction between the cognate effect and frequency effect, both present and statistically-significant, is likely because these effects are determined at different stages of the model and their connections are not set to have a large amount of interaction. Empirically, this interaction is approximately 1,000x less powerful than either effect alone, yet still determinable. When all variables are included into a statistical model, approximately 46% of the effects (*adjusted-*R^2 = 0.46, see section 7.5) are accounted for by the current data.

Jacobs and Grainger (1994) sets forth a guide for model evaluation, and applies these to 15 early (pre-) computational (or "algorithmic") models of visual word recognition. Models are classified according to sets of properties, separated by: fundamental recognition process (family); design expliciteness (format); featured test paradigms (task); output dependent variable; "simplicity", comprising 8 binary absent or present subfeatures, deterministic\probabilistic output, localist\distributed representation, macroscopic\microscopic performance prediction; modular\interactive feedback for representation-levels, ordinal\interval scale predictions, performance\learning algorithm, serial\parallel search & verification mechanism, static\dynamic accumulative processing functions; and testable effects. Following this, a proposal for "standards of model evaluation": *descriptive adequacy*, corresponding to format, dependent variable, static\dynamic, and ordinal\interval features; *generality*, corresponding to the features of task, dependent variable, and effects; *simplicity\falsifiability*, which approximately corresponds to the number of hypotheses, representation levels, or diagrammatic interconnections employed by the model, weighted by the amount of explainable phenomenon that each additional parameter accounts for; *explanatory adequacy*, referring to the "assumptions" built into the model, ad-hoc[29] or empirical; and lastly, *modifiability, research generativity, equivalence class, model-completeness,* and *(neurobiological)*

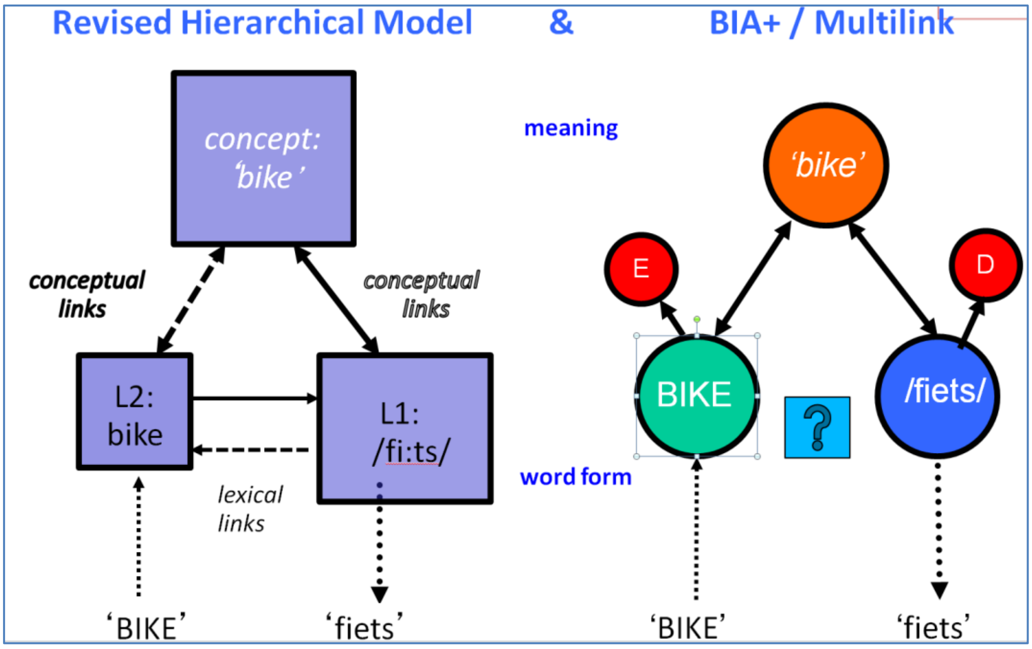
[29] "[. . .] ad hoc assumption are used here as synonyms for a hypothesis or algorithm that serves to explain no more effects than the one it was introduced to explain. It should be useful to note that the fact that a model includes an ad hoc assumption does not imply that the model is useless." (Jacobs & Grainger, 1994: 1319)

plausibility are added, but left unspecified due to space considerations (Jacobs & Grainger, 1994: 1312-1320). These same criteria are applied to Multilink, RHM, the BIA+ in table 21 (below):

	RHM	BIA+	Multilink
Family	Dual-route	Interactive activation	Interactive activation
Format	Descriptive	Computational	Computational
Task	Word-naming, lexical decision, translation	Cognate Recognition, Lexical Decision	Cognate Recognition, Language Decision, Lexical Decision, Translation
DV	RT	RT, accuracy	RT, accuracy
deterministic\probabilistic	Deterministic	Deterministic	Deterministic
localist\distributed	Localist	Localist	Localist
macroscopic\microscopic	Macroscopic	Microscopic	Microscopic
modular\interactive	Modular	Interactive activation	Interactive activation
ordinal\interval	Ordinal	Interval	Interval
performance\learning	Performance	Performance	Performance
serial\parallel	Serial	Parallel	Parallel
static\dynamic	Static	Dynamic	Dynamic
testable effects	Bilingual asymmetry, language proficiency	Cognate, frequency, length	Bilingual asymmetry, language proficiency, cognate, frequency, length

Table 21. Property Overview of the RHM, BIA+, and Multilink models, á la Jacobs & Grainger (1994: 1313, Table 1). References and details for each model found in section 5.1

Figure 24. Architecture Comparison Of the RHM & BIA+/Multilink models. Note that Multilink does not hypothesize word association.



It is difficult to compare the three models, although they have several commonalities. First, the RHM is pre-computational. Second, the RHM and BIA+ both have limited task specifications and testable effects. Matched to Multilink, a computational model with a greater number of task specifications and testable effects, comparing the three models hardly seems apt. Third, the temporal-informational factor: RHM and BIA+ models debuted 20 and 12 years ago, respectively. Although both have continued to inform and find use in recent literature, the landscape of psycholinguistic literature has changed in the intervening years. What was theoretically-informative for model-building at that time is not necessarily as informative as it was then, and factors that were not considered relevant or strong at that time might now be considered important. In terms of model ranking, Multilink wins the race by a landslide: a computational model outputting interval-scale data offers greater descriptive adequacy than any verbal model can compete with, though Multilink is equal with the BIA+ in this case. Multilink has greater generality than both BIA+ and RHM, having a greater number of testable effects, and task simulations. Multilink is absolutely a more complex model than the BIA+ and the RHM, but its complexity is a trade-off for greater theoretical coverage. Rating each model in terms of explanatory-adequacy, the number of ad-hoc vs empirical assumptions, is more difficult: each seems approximately equal in this instance. Multilink wins this race purely because it has greater specification and theoretical coverage. It is currently not clear how Multilink would fare in a similar evaluation against other, equally-advanced monolingual & bilingual computational or mathematical models of visual word recognition, such as WEAVER++ (Roelofs, 1997; Roelofs et al., 2013), SOLAR (Davis, 2001), CDP+ (Perry, Ziegler, & Zorzi, 2007), Bayesian Reader (Norris, 2006), or the DRC (Coltheart, et al. 2001).

Referring back to section 6.1.i-iii, the research questions, we observe that Multilink does not induce a translation direction effect, but it does induce a cognate effect, and a frequency effect, fulfilling 2 of the 3 research questions.

8.2 **Facilitation Effects**

What does Multilink need to improve? Considering only the main effects — direction, cognate, and frequency — and their interactions, two actions could greatly improve the output of the model: parameter adjustment, and effect inclusion.

Forward translation direction is experimentally-observed to have a facilitatory effect on word-naming latency. At this time, Multilink does not replicate this effect, and might actually trend towards a backwards translation direction facilitation effect. Its disclusion is understandable, as the nature of translation direction facilitation has been under debate for several years, but the debate has been weighted towards the forward facilitation effect by the results of this study, and Pruijn (2015). This could be included by adjusting the weight of the language nodes, to reflect a higher resting activation level for the L2 production node, when activated by the L1 language node.

The cognate effect is already present, a recognizably-strong effect. However, it is more powerful in the model than experimental data suggests that it needs to be. The candidate-ranking algorithm would seem to bear much of the blame here, given that it ranks outputs primarily on their status as cognates or noncognates. Multilink's cognate effect size should be reduced. This could perhaps be done by reducing the input of LD ranking into the candidate selection, or by reducing the activation spread of the orthographic-phonological or orthographic-lexical representation connections, which should demote the role of orthography in determining phonology and the final output.

The influence of the frequency facilitation effect within the ranking algorithm needs to be embiggened. Frequency facilitation is an experimentally-strong effect, nearly equal to the cognate effect, but is not being replicated as such in the output. To affect this, the candidate-ranking algorithm should be changed to increase the input from frequency data already included in the lexicon, or the activation-input of frequency into the phonology representation should be boosted.

It is possible that when the main effects are adjusted that the significant interactional effects — forward translation direction effect & cognate effect interaction, and the cognate effect & frequency effect interaction — will be observed. However, if this is not the case (which seems likely, since parameters in a computational model generally need to be hard-set), the orthographic-language activation input multiplier should be adjusted to reflect a strength-level associated with the empirical data. Current results recommend that this connection should be increased by a factor of 2.3x. For the interaction of cognate effect & frequency effect, the LD candidate-ranking or orthographic-phonological input and its connection to the phonology representation should be augmented by a factor of 1.6x to adequately replicate the interaction between cognates and frequency.

8.3 **Latent Phenomena**

The recommended improvements only cover the first action: parameter adjustment, and only include the effects that have received wide experimental interest. What about the latent effects observed in regression modelling (section 7.5)? Empirically, phonetic onset was found to have a significant effect; and Multilink was found to process cognates of varying length in a significant, and non-empirically-observed, fashion.

Summing up the interaction of the cognate effect and word-length, it is currently unknown why the model outputs with regards to word-length in this fashion, steadily lowering the mean cycle-time as word length increases. However, it could be a confound: when a word form is input, it activates all orthographic neighbours, shorter or longer words receiving reduced activation. However, the lexicon varies word-length between 3-8 letters; selection is faster for lengthier words because the orthographic neighbourhood is less dense after 6-letter words (the mean of Multilink's lexicon (Dijkstra, 2010: 409), leading to less competition between word forms to rank as output, and thus causing faster activation. Whatever the cause of this model-facilitatory effect of stimulus length on cognates in the model, it does not have an empirical reflection. Experiment data thus far submits that cognate status does not interact significantly with word-length. Referring back to section 7.5.4 and figure 23 (page 59), cognates of 4-8 letters in length hover between 750-800 ms. 3-letter cognates are the exception, with a latency at ~820 ms. Noncognates show a "hill" type curve, with shorter lengths being processed faster, and slowly rising before declining again. Multilink can take one of two positions to rectify this: the conservative position of "length effect", as demonstrated by New, et al. (2006) and O'Regan & Jacobs (1992), finding that longer word length is inhibitory (Multilink should show an incline for the output curve with respect to length). Nonetheless, the notion of the "word length effect" is debated, and evidence has also shown a null effect: another position would be to disregard word length from output ranking, so that it has no effect on ultimate output cycle-time, if that is possible.

Phonetic onset (referring back to sections 5.2.2.4, and 7.5.4) is a trickier notion to discuss for the model, due to phonetic/phonological/technical variables. Voiced and voiceless consonants should exhibit equal latency, but the latency of voiceless onsets is nearly 250 ms higher than that of voiced onsets. This issue was explored by Rastle & Davis (2002), ultimately concluding that matching for onset phoneme (as was partially done by Pruijn (2015) and the current study) is an inadequate control method. But this effect can be controlled by matching the conditions of the syllabic onset (for instance, if the onset is complex, ie [spl] in the word "splinter"). Their results predominantly recommend a switch to more accurate measures, such as hand-coding or algorithmic-coding of acoustic waveform data, or more precise phonation collection methods[30]. Technical issues aside, this still does not explain why there would be a directional interaction with phonetic onset. Recalling Figure 22 (page 58), tokens with a vocalic onset are produced slower in the forward condition than vocalic onsets in the backward condition. Due to the extremely small vocalic onset sample size, 15 total (approximately 7 in each direction), discerning a cause or correlation is statistically-improbable; there is simply not enough power. Three possible correlating factors

[30] This would have to target a different, or multiple, articulatory mechanism(s); the voice-key, aptly, distinguishes primarily on the phonological feature of voicing.

stand out: first, a *syllable frequency factor*, *articulation time*, and the phonological rule of "*maximal onset*"[31].

Syllable frequency is, simply, the frequency that each syllable occurs in the lexicon. Cholin, Levelt, & Schiller (2006) conducted a series of experiments demonstrating that speakers use and access a mental syllabary, and that this syllabary is especially sensitive to first-syllable frequency manipulations, and Carreiras, Alvarez, & De Vega (1993) found evidence suggesting that the "syllable frequency effect" is not simply due to the frequency of phoneme co-occurence (which lends additional credence to the concept of an online mental syllabary, contradicting Seidenberg's (1987, 1989) Orthographic Redundance hypothesis). Furthermore, Perea & Carreiras (1998) show that syllables have their own neighbourhood effect, with high-frequency syllables engendering an inhibitory effect. Barber, Vergara, & Carreiras (2004), and Carreiras, Mechelli, & Price (2006) demonstrate with electrophysiological and hemodynamic experiments, respectively, that the brain reacts inversely towards high lexical frequency, and high initial syllable frequency: the former is facilitative, and the latter is inhibitive. Reicker, et al. (2008) found that complex onsets create a significant effect on speech motor control and planning using hemodynamic measurements in a lexical decision task. Lastly, Conrad & Jacobs (2004) demonstrate, using the Functional Units Model (Rey, 1998; Richter, 1999) to illustrate their argument, that future computational models of visual word recognition cannot ignore and disclude the stand-alone syllable frequency effect.

Articulation time is the amount of time that a human being requires to place speech organs, articulators, into position, commence phonation, and for the sound to exit the lips. Palo et al. (2015), Rastle et al. (2005), and Kawamoto et al. (2008) provide evidence that there is a significant delay for some types of consonants due to articulatory properties, and that there is a difference between articulation time, when phonation commences, and acoustic latency, when phonation is experimentally-measured. Thus far, it is not a huge confound for word-naming studies, but the increasing need for accuracy will force future studies to consider options for increasing experimental control and statistical power.

Maximal onset, first described by Kahn (1976), is a general cross-linguistic phonological principle stating that, for languages that allow onsets (ie, English and Dutch), "[. . .] given a choice between affiliating a consonant to a coda or to an onset, affiliating with the onset is preferable [. . .]" (Hulden, 2006: 90). That is, if a language allows onsets, it will almost always prefer to place a consonant into the onset position within the syllable. Initial syllables, as evidenced above, occupy a privileged place in the word; it might then be

[31] Acknowledgements to Dr. Carlos Gussenhoven, and Dr. Francisco Torreira for sharing references and information concerning the maximal onset principal, and syllable frequency effects.

rationally assumed that the initial onset occupies a privileged position within the word as well. For words that do not have a consonant in the onset position, this would be considered a syllable constraint violation (in Optimality Theory terms) if the onset was vocalic. Hypothetically, constraint violation could lead to higher naming latencies, particularly for the initial syllable. Constraint violation might also interact with syllable frequency.

Whatever the cause(s) or ultimate correlating factor(s), previous research clearly demonstrates that Multilink must include a level of syllabic representation, and also a lexicon of syllable frequencies, similar to the syllabary built into WEAVER++. Perhaps, a second experiment can be conducted to determine the strength of a translation direction and syllabic feature interaction[32].

8.4 **Directions For Future Research**

9 important potential future directions for Multilink have been noted per the results of the current study, presented in no particular order.

i. *Increasing typological distance* — as stated in section 7.6.3, Multilink, as a lexical information-processing system, is implicitly optimized towards Dutch & English, (West) Germanic languages, and the Indo-European language family more broadly. Slowly broadening the geographic distance between languages and language families is one starting method. A future study could test Multilink first with a North-Germanic language, such as Swedish, then with a Romance or Slavic Indo-European language, such as French or Czech, and could end with any unrelated language (an isolate, perhaps, such as Basque) that has reliable and recent lexical frequency data tables available (Japanese, with the BCCWJ[33] (Maekawa et al., 2014), courtesy of NINJAL, is a good first candidate for the "maximum-distance" comparison). This would also help test assumptions about how typologically-distant cognate-loans might be recognized and produced. As typological distance expands, the potential for script expansion also presents itself: Multilink is capable of modelling only languages with the Latin script. A first candidate would need to have comparable properties to the latin script: left-to-right reading direction, phonetic representation (as opposed to an abjad or alphasyllabary), vowels represented, etc. The Cyrillic script, used in Russia and several states of Eastern Europe, is a potential first test option.

[32] For an overview concerning variation in syllable structure and phonotactic properties of syllables, the reader is referred to Duanmu (2008), *Syllable Structure: The Limits Of Variation*, and Cairns & Raimy (2011), *Handbook Of The Syllable*.

[33] http://www.ninjal.ac.jp/english/products/bccwj/

ii. *Deprecate the length restriction* — Multilink is currently limited to 8-9 letters for its maximum length. While double that of the BIA+, this is still a very limiting restriction. If the restriction were deprecated, then the addition of a variable inhibitory and facilitatory word length effect, such as seen in New, et al. (2006), could be introduced to Multilink. This could be supported with a second bilingual word-naming experiment using stimuli that specifically manipulate length and control for the other factors. Either way, it is clear that Multilink needs correction in its length-parameter.

iii. *Inclusion Of Syllabic Properties* — as stated above in section 8.3, the initial syllable is known to generate special effects that affect the entire word form. Accounting for only the phonetic onset type or the onset complexity in the final dataset is not enough. A supplemental model "mental" syllabary would appear to be required. In addendum to this, a new experiment employing a more precise word onset collection mechanism might be required as well, in order to counteract the voice-key onset delay, noted to be the single largest effect of this study. Per the suggestions of Rastle & Davis (2002), an acoustic waveform algorithmic onset detector is the most convenient solution.

iv. *Parameter Adjustment* — the 5 main independent variable effects observed in the experimental data need to be adjusted within the model to better represent their empirical task counterparts.

v. *Integrated Scaling Models* — per the recommendation of Jacobs & Grainger (1994), outputting model data in purely cycle-time is unnecessarily opaque. The goal of any predictive scientific model, computational or pre-computational, should be to clarify the processi and consequences of the hypotheses being tested. In the interest of expedited interpretation, Multilink should have an integrated statistical model that uses acknowledged and valid empirical measurements to predict latency data from output cycle-time (cycle-time would still be an output component). This would first require the aforementioned alterations to be accomplished, and second, for a set varied of statistical models to be tested against new experimental data to determine which scaling algorithm results in the best empirical fit. This would not be a trivial task, but would potentially help Multilink be a useful tool to a larger audience of researchers who do not have the skills to create or interpret complex statistical models.

vi. *Increased Linguistic Scope* — Multilink is a model of word-naming, concerned with recognizing and producing isolated words in the lexicon in one of (presently) two languages. As remarked in section 6.3.iv, this is an adequate starting point, but speech, bilingual or monolingual, does not use single words; it is, in fact, much more complex. Muysken (2005: 47) produces a hierarchical table (reproduced on the next page, Figure 24) showing how each "layer" of language multiplies the number of units in speech, and thus multiplies the complexity as well.

The different levels of analysis assumed by linguists

Text	unit of a single communicative event
Turn	unit of the part taken by a single speaker during that event
utterance or sentence	all directly connected uttered material within a turn
clause	unit describing a single state of affairs or proposition
phrase	unit within a clause consisting of several closely linked words
word	minimal unit that can be uttered separately
morpheme	minimal unit that has meaning
phoneme	minimal sound component
feature	minimal sound specification

Figure 25. "The different levels of analysis assumed by linguists" (Muysken, 2005: 47).

At present, according to this analysis, Multilink's representations range from the individual feature-level, to the word-level, but not beyond[34]. Van Assche et al. (2012) reviews studies of bilingual lexical access in a sentence context, situating findings according to the BIA+ model. The cognate effect was present in reviewed studies. In light of this evidence, it is inevitable that Multilink will need to increase the scope of its use, first to the phrase level, then clause, and finally up to sentence level. Turn and text levels could theoretically be reached once model sentences can be accurately output. In line with increasing the scope of complexity, Multilink will also need to be able to model bilingual lexical processing for other lexical categories, such as verbs, or adjectives. In particular, it would be noteworthy to experimentally check for differences in cognitive patterns of speakers from languages which do not distinguish adjectives from verbs, such as Okinawan Ryukyuan (Japonic family language spoken in Okinawa, Japan), or Montagnais (Algic-Algonquian family language of Northeastern Canada) (Hammarström, et al. 2015).

vii. *Built-in Lexicon Adjustment* — This would be a marginal improvement, but still useful. Multilink needs to be accessible in order to reach a wide audience. There needs to be built-in programming for users to easily add, remove, and modify the lexicon, and word association files. One example: if a future study wishes to model L2 acquisition, frequency data in Multilink's lexicon needs to be hand-altered. This is impractical, and inconvenient. Programming frequency alteration into Multilink's user-interface would expedite this process. It should be possible to add completely new tokens into the lexicon just as easily. Perhaps,

[34] Technically, Multilink skips the "morpheme" level, since there are no decomposable words in Multilink's lexicon

in the future, the representation activation levels themselves could be input-adjusted by a numerical factor too.

viii. *Learning Algorithm* — Currently, Multilink uses a "performance" algorithm (as seen in Table 21, section 8.1, page 70). Jacquet & French (2002) suggested, for the BIA+ model, adopting a statistical-learning algorithm. This would allow Multilink to require less hand-coding of information, and be able "experience" the language by having the information input as training data. Most importantly, this would potentially allow the model to develop an association-index and a lexicon on its own, thus negating the potential for the type of errors seen in Table 6, section 6.4 (page 35) (Jacquet & French, 2002: 204-205). Of all the potential upgrades listed so far, this would by far be the most beneficial.

ix. *Concreteness Scores* — The addition of concreteness scores into the model architecture is empirically-required. Kroll & Merves (1986), Barber et al. (2013), and Jessen et al. (2000) have shown that there is a cognitive disparity for the processing of concrete and abstract words, resulting in concrete words being accessed and produced at a faster rate. Van Hell & De Groot (1998a, b) show evidence of bilingual concreteness facilitation effect in both nominals and verbals. Concreteness even appears to interact with cognate status, and lexical category, but this could also be confound with "context availability", from the hypothesis of the same name, claiming that concreteness effects arise from variations in contextual information — prior knowledge — that makes concrete words faster to retrieve than abstract words (objects conceptually-associated with concrete words are encountered more often than objects conceptually-associated with abstract words). At this time, there are no lexical cognitive processing models that acquire and employ concreteness scores within the architecture. Multilink is in a prime position to fill this deficiency.

x. *Up-to-date frequency data* — A computational model is only as good as the data that informs it. The CELEX (Baayen, Piepenbrock, & Van Rijn, 1993) corpus is used by Multilink for its English & Dutch frequency data. In the interest of accurate replication, this frequency data is considered subpar, and recent replacements are available. SUBTLEX-US (Brysbaert & New, 2009) and SUBTLEX-NL (Keuleers, Brysbaert, & New 2010) are recommended for this (these were additionally used by Pruijn (2015)), and future work on Multilink should first look to replace the current frequency data with a newer equivalent, as long as the corpus is considered accurate and ecologically-valid.

9 Conclusion

This study has examined the performance of the Multilink model of bilingual word recognition and production of Dijkstra and Rekké (2010), and compared the performance of the model to the recent empirical study (Pruijn, 2015, in collaboration with Peacock); this study used a precisely-balanced stimulus set in a translation production task to clarify the ongoing debate between the various models of bilingual lexical processing and access, such as the RHM of Kroll & Stewart (1994, 2010), concerning the language direction facilitation effect: is L1→L2 (forward) production processed faster than L2→L1 (backward) production? By dividing the stimulus into 8 distinct categories, each manipulating the 3 binary variables of translation direction, cognate status, and frequency, orthogonally, significant effects and interactional effects can be observed. The results of Pruijn (2015) clearly show an overall forward facilitation effect — rejecting the access routes proposed by the RHM — and a cognate facilitation effect, a frequency facilitation effect, and two significant interactional facilitation effects: direction & cognate status, and cognate status & frequency.

Utilizing exactly the same stimulus set, with its orthogonal manipulations, as input to Multilink, this model's output, measured in "cycle-times", has been compared to the empirical findings of Pruijn (2015), and analyzed via four statistical tests: Spearman's rank correlation, analysis of variance, generalized regression modelling, and statistical divergence testing. Given the high accuracy of the output (98.44%), accuracy data was not included in the testing. It was inferred that the model performed well, not only in terms of accurate output, but also as a computational replication of bilingual lexical access (as divined from latency measurements in naming tasks). However a comparison of plotted means (Figure 13 & 14, page 42) showed that Multilink is not replicating some empirically-observed effects.

Correlational results at the item level were, unexpectedly, largely non-significant for each of the 8 categories, even between latency and frequency; cycle-time & latency did show correlations within the larger forward, backward, and global conditions. Correlations did show a pattern of larger correlation coefficients between cycle-time & frequency for non-cognates, Levenshtein Distance (LD) has a stronger pattern of correlations with cognate conditions, and frequency strongly correlates with cycle-time in noncognate conditions. Few other correlational patterns were perceived. A final mean-regressed correlation shows that cycle-time, on average, strongly correlates with latency and LD.

ANOVA testing of the empirical and model data shows several levels of effects present: latency predicts significant effects with translation direction, cognate status, frequency, and marginal interactions between translation direction & cognate status, and cognate status & frequency (these were measured as significant by Pruijn (2015), see section 7.4.1 footnote 13, page 49). Cycle-time, conversely, only predicts significant effects with cognate status and frequency. Effect size measurements were compared, deducing that Multilink is overrating the cognate effect, underscoring the frequency effect, and is largely missing the translation direction effect and interactional effects.

Regression analysis follows, using a generalized additive model (GAM) to test for interactions between the dependent variables, independent variables, and the controlled variables, and their two-way interactions. This helped assess whether controlled variables were adding noise into the data, and the total amount of deviance explained by the current data. Of the controlled variables, phonetic onset was predicted to interact significantly with latency and assigned the largest amount of deviance within the empirical data. Model data predicts sole interaction between cognate status and word-length, which revealed that Multilink deviates from the empirical results with respect to word-length (Figure 23, page 60). However, none of these effects contributed so much noise to the data that the experimental effects were unobservable. Possible reasons for these interactions were discussed in section 8.3 (page 72-74).

Finally, the distributional divergence was measured, using a model-approximate χ^2 test, Kolmogorov-Smirnov test, and a Kullback-Leibler Divergence test. These tests measure how close the datapoint distributions (as probability distributions or cumulative distributions), match each other, and the output statistic measures the degree of this divergence in Shannons/bits. Results of these tests showed that Multilink, as a model, actually fits the data quite well. Despite deficiencies such as the overpowered cognate facilitation and the underpowered frequency facilitation effect, totalled over all datapoints, the distributions match closely, with a measured maximum divergence of less than a single bit.

These results were then discussed in the context of other models of the bilingual lexicon, and possibilites for future studies and improvements to Multilink were planned in section 8.4. Some of these suggestions include: changing to a statistical-learning algorithm, addition of syllabic representation into the model, and using more typologically-distant languages for further modelling (once the necessary adjustments are made).

Given the literature previously reviewed, it is clear that this computational model does not include all the factors that are necessary to explain multilingual lexical access. But with the factors that it does account for Multilink performs admirably, and holds great promise for the future of lexical access studies, and psycholinguistics as a whole. Based on the recent empirical results, several changes should be made: the ranking formula, which determines the output candidates, needs to be updated, reflecting an observable forward translation direction facilitation effect, LD must be balanced with respect to the output, the frequency facilitation effect must be strengthened, and the smaller interactional effects between cognate status & frequency and translation direction & cognate status should become part of the ranking algorithm. Once this is accomplished, Multilink can be extended and enhanced even further, upgrading it into a versatile and valuable tool for future linguistic and cognitive studies.

10 References

Abutalebi, J. (2008). Neural aspects of second language representation and language control. Acta Psychologica, 128(3), 466-478. http://dx.doi.org/10.1016/j.actpsy.2008.03.014

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. Journal Of Neurolinguistics, 20(3), 242-275. http://dx.doi.org/10.1016/j.jneuroling.2006.10.003

Altarriba, J., Bauer, L., & Benvenuto, C. (1999). Concreteness, context availability, and imageability ratings and word associations for abstract, concrete, and emotion words. Behavior Research Methods, Instruments, & Computers, 31(4), 578-602. http://dx.doi.org/10.3758/bf03200738

Asaip.psu.edu,. (2015). Beware the Kolmogorov-Smirnov test! — Astrostatistics and Astroinformatics Portal. Retrieved 17 December 2015, from https://asaip.psu.edu/Articles/beware-the-kolmogorov-smirnov-test

Baayen, R. H., Piepenbrock, R., & van H, R. (1993). The {CELEX} lexical data base on {CD-ROM}.

Bachoud-Lévi, A., Dupoux, E., Cohen, L., & Mehler, J. (1998). Where Is the Length Effect? A Cross-Linguistic Study of Speech Production. Journal Of Memory And Language, 39(3), 331-346. http://dx.doi.org/10.1006/jmla.1998.2572

Baddeley, A., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. Journal Of Verbal Learning And Verbal Behavior, 14(6), 575-589. http://dx.doi.org/10.1016/s0022-5371(75)80045-4

Barber, H., Otten, L., Kousta, S., & Vigliocco, G. (2013). Concreteness in word processing: ERP and behavioral effects in a lexical decision task. Brain And Language, 125(1), 47-53. http://dx.doi.org/10.1016/j.bandl.2013.01.005

Barber, H., Vergara, M., & Carreiras, M. (2004). Syllable-frequency effects in visual word recognition: evidence from ERPs. Neuroreport, 15(3), 545-548.

Bassetti, B. (2013). Bilingualism and writing systems. The Handbook of Bilingualism and Multilingualism, Second Edition, 649-670.

Bhatia, T., & Ritchie, W. (2012). The handbook of bilingualism and multilingualism. Chichester, West Sussex, UK: Wiley-Blackwell.

BRYSBAERT, M., & DUYCK, W. (2010). Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after fifteen years of service?. Bilingualism, 13(03), 359-371. http://dx.doi.org/10.1017/s1366728909990344

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. Behavior Research Methods, 41(4), 977-990. http://dx.doi.org/10.3758/brm.41.4.977

Brysbaert, M., Stevens, M., De Deyne, S., Voorspoels, W., & Storms, G. (2014). Norms of age of acquisition and concreteness for 30,000 Dutch words. Acta Psychologica, 150, 80-84. http://dx.doi.org/10.1016/j.actpsy.2014.04.010

Brysbaert, M., Warriner, A., & Kuperman, V. (2013). Concreteness ratings for 40 thousand generally known English word lemmas. Behav Res, 46(3), 904-911. http://dx.doi.org/10.3758/s13428-013-0403-5

Bultena, S. (2013). Bilingual processing of cognates and language switches in sentence context (Doctoral Dissertation). Radboud Universiteit Nijmegen.

Carreiras, M., Alvarez, C. J., & Devega, M. (1993). Syllable frequency and visual word recognition in Spanish. Journal of memory and language, 32(6), 766-780.

Carreiras, M., Mechelli, A., & Price, C. J. (2006). Effect of word and syllable frequency on activation during lexical decision and reading aloud. Human brain mapping, 27(12), 963-972.

Center For Reading Studies,. (1979). PROCESSING FACILITATION IN A LEXICAL DECISION TASK. University Of Illinois Urbana-Champaign.

CHINCOTTA, D., & UNDERWOOD, G. (1998). Non temporal determinants of bilingual memory capacity: The role of long-term representations and fluency. Bilingualism, 1(02), 117-130. http://dx.doi.org/10.1017/s1366728998000248

Cholin, J., Levelt, W. J., & Schiller, N. O. (2006). Effects of syllable frequency in speech production. Cognition, 99(2), 205-235.

Christiansen, M. H., & Chater, N. (Eds.). (2001). Connectionist psycholinguistics. Greenwood Publishing Group.

CHRISTOFFELS, I., DEGROOT, A., & KROLL, J. (2006). Memory and language skills in simultaneous interpreters: The role of expertise and language proficiency. Journal Of Memory And Language, 54(3), 324-345. http://dx.doi.org/10.1016/j.jml.2005.12.004

Cicek, E., & Yilmaz, A. (2013). A new Morse code scheme optimized according to the statistical properties of Turkish. Turkish Journal Of Electrical Engineering & Computer Sciences, (21), 804-811. http://dx.doi.org/doi:10.3906/elk-1110-11

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. Psychological Review, 100(4), 589-608. http://dx.doi.org/10.1037//0033-295x.100.4.589

Coltheart, M., Davelaar, E., Jonasson, J., & Besner, D. (1977). Access to the Internal Lexicon, 535-555.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual-route cascaded model of visual word recognition and reading aloud. Psychological Review, 108(1), 204-256. http://dx.doi.org/10.1037//0033-295x.108.1.204

Conrad, M., & Jacobs, A. (2004). Replicating syllable frequency effects in Spanish in German: One more challenge to computational models of visual word recognition. Language and Cognitive Processes, 19(3), 369-390.

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. Journal Of Memory And Language, 50(4), 491-511. http://dx.doi.org/10.1016/j.jml.2004.02.002

Dahan, D., Magnuson, J., & Tanenhaus, M. (2001). Time Course of Frequency Effects in Spoken-Word Recognition: Evidence from Eye Movements. Cognitive Psychology, 42(4), 317-367. http://dx.doi.org/10.1006/cogp.2001.0750

Davelaar, E., Coltheart, M., Besner, D., & Jonasson, J. (1978). Phonological recoding and lexical access. Memory & Cognition, 6(4), 391-402. http://dx.doi.org/10.3758/bf03197471

Davis, C. J. (2001). The self-organising lexical acquisition and recognition (SOLAR) model of visual word recognition (Doctoral dissertation, ProQuest Information & Learning).

de Bot, K. (2004). The Multilingual Lexicon: Modelling Selection and Control. International Journal Of Multilingualism, 1(1), 17-32. http://dx.doi.org/10.1080/14790710408668176

de Bot, K., Cox, A., Ralston, S., Schaufeli, A., & Weltens, B. (1995). Lexical processing in bilinguals. Second Language Research, 11(1), 1-19. http://dx.doi.org/10.1177/026765839501100101

Degroot, A., Dannenburg, L., & Vanhell, J. (1994). Forward and Backward Word Translation by Bilinguals. Journal Of Memory And Language, 33(5), 600-629. http://dx.doi.org/10.1006/jmla.1994.1029

Dijkstra, T. (1990). Cross-modal contacts between graphemes and phonemes: explorations in bimodal processing (Doctoral Dissertation). Katholieke Universiteit Te Nijmegen.

Dijkstra, T. (2005). Bilingual Visual Word Rec

Dijkstra, T., & Rekké, S. (2010). Towards a localist-connectionist model of word translation. ML, 5(3), 401-420. http://dx.doi.org/10.1075/ml.5.3.08dij

Dijkstra, T., & van Heuven, W. (2002). The architecture of the bilingual word recognition system: From identification to decision. Bilingualism, 5(03). http://dx.doi.org/10.1017/s1366728902003012

Dijkstra, T., & Van Heuven, W. J. (1998). The BIA model and bilingual word recognition. Localist connectionist approaches to human cognition, 189-225.

Dijkstra, T., Lormans, E., Rekké, S., & Roelofs, A. Multilink: a computational model for word translation.

Dijkstra, T., Moscoso del Prado Martín, F., Schulpen, B., Schreuder, R., & Harald Baayen, R. (2005). A roommate in cream: Morphological family size effects on interlingual homograph recognition. Language And Cognitive Processes, 20(1-2), 7-41. http://dx.doi.org/10.1080/01690960444000124

Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. Journal Of Experimental Psychology: Learning, Memory, And Cognition, 33(4), 663-679. http://dx.doi.org/10.1037/0278-7393.33.4.663

Edwards, J., Pexman, P., Goodyear, B., & Chambers, C. (2005). An fMRI investigation of strategies for word recognition. Cognitive Brain Research, 24(3), 648-662. http://dx.doi.org/10.1016/j.cogbrainres.2005.03.016

Feustel, T., Shiffrin, R., & Salasoo, A. (1983). Episodic and lexical contributions to the repetition effect in word identification. Journal Of Experimental Psychology: General, 112(3), 309-346. http://dx.doi.org/10.1037/0096-3445.112.3.309

Fiebach, C., Friederici, A., Müller, K., & Cramon, D. (2002). fMRI Evidence for Dual-routes to the Mental Lexicon in Visual Word Recognition. Journal Of Cognitive Neuroscience, 14(1), 11-23. http://dx.doi.org/10.1162/089892902317205285

French, R., & Jacquet, M. (2004). Understanding bilingual memory: models and data. Trends In Cognitive Sciences, 8(2), 87-93. http://dx.doi.org/10.1016/j.tics.2003.12.011

Frost, R. (1995). Phonological computation and missing vowels: Mapping lexical involvement in reading. Journal Of Experimental Psychology: Learning, Memory, And Cognition, 21(2), 398-408. http://dx.doi.org/10.1037//0278-7393.21.2.398

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. Psychological Bulletin, 123(1), 71-99. http://dx.doi.org/10.1037//0033-2909.123.1.71

G. van Hell, J., & de Groot, A. (1998). Disentangling Context Availability and Concreteness in Lexical Decision and W ord Translation. The Quarterly Journal Of Experimental Psychology Section A, 51(1), 41-63. http://dx.doi.org/10.1080/713755752

Gee, N., Nelson, D., & Krawczyk, D. (1999). Is the Concreteness Effect a Result of Underlying Network Interconnectivity?. Journal Of Memory And Language, 40(4), 479-497. http://dx.doi.org/10.1006/jmla.1998.2627

Gooskens, C., & Heeringa, W. (2004). Perceptive evaluation of Levenshtein dialect distance measurements using Norwegian dialect data. Lang. Var. Change, 16(03). http://dx.doi.org/10.1017/s0954394504163023

Gow Jr, D. W., Melvold, J., & Manuel, S. (1996, October). How word onsets drive lexical access and segmentation: Evidence from acoustics, phonology and processing. In Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on (Vol. 1, pp. 66-69). IEEE.

Grainger, J., & Jacobs, A. M. (1998). Localist connectionist approaches to human cognition. Psychology Press.

Grainger, J., Muneaux, M., Farioli, F., & Ziegler, J. (2005). Effects of phonological and orthographic neighbourhood density interact in visual word recognition. The Quarterly Journal Of Experimental Psychology Section A, 58(6), 981-998. http://dx.doi.org/10.1080/02724980443000386

GREEN, D. (1986). Control, activation, and resource: A framework and a model for the control of speech in bilinguals. Brain And Language, 27(2), 210-223. http://dx.doi.org/10.1016/0093-934x(86)90016-7

GREEN, D. (1998). Mental control of the bilingual lexico-semantic system. Bilingualism, 1(02), 67. http://dx.doi.org/10.1017/s1366728998000133

Griffin, Z., & Bock, K. (1998). Constraint, Word Frequency, and the Relationship between Lexical Processing Levels in Spoken Word Production. Journal Of Memory And Language, 38(3), 313-338. http://dx.doi.org/10.1006/jmla.1997.2547

GROSJEAN, F. (1998). Studying bilinguals: Methodological and conceptual issues. Bilingualism, 1(2), 131-149. http://dx.doi.org/10.1017/s136672899800025x

Hammarström, Harald & Forkel, Robert & Haspelmath, Martin & Bank, Sebastian. 2015. Glottolog 2.6. Jena: Max Planck Institute for the Science of Human History. (Available online at http://glottolog.org, Accessed on 2015-12-26.)

Heredia, R. (1995). CONCRETENESS EFFECTS IN HIGH FREQUENCY WORDS: A TEST OF THE REVISED HIERARCHICAL AND THE MIXED MODELS OF BILINGUAL MEMORY REPRESENTATIONS (Doctoral Dissertation). University of California At Santa Cruz.

Heredia, R. (1997). Bilingual Memory and Hierarchical Models: A Case for Language Dominance. Current Directions In Psychological Science, 6(2), 34-39. http://dx.doi.org/10.1111/1467-8721.ep11512617

Hulden, M. (2006). Finite-state syllabification. In Finite-State Methods and Natural Language Processing (pp. 86-96). Springer Berlin Heidelberg.

Indiana University,. (2007). Frequency and the Emergence of Prefabs: Evidence from Monitoring1. Indiana University.

Jacobs, A., & Grainger, J. (1994). Models of visual word recognition: Sampling the state of the art. Journal Of Experimental Psychology: Human Perception And Performance, 20(6), 1311-1334. http://dx.doi.org/10.1037//0096-1523.20.6.1311

Jared, D., & Szucs, C. (2002). Phonological activation in bilinguals: Evidence from interlingual homograph naming. Bilingualism, 5(03). http://dx.doi.org/10.1017/s1366728902003024

Jessen, F., Heun, R., Erb, M., Granath, D., Klose, U., Papassotiropoulos, A., & Grodd, W. (2000). The Concreteness Effect: Evidence for Dual Coding and Context Availability. Brain And Language, 74(1), 103-112. http://dx.doi.org/10.1006/brln.2000.2340

Kahn, D. (1976). Syllable-based generalizations in English phonology (Vol. 156). Bloomington: Indiana University Linguistics Club.

Kawamoto, A. H., Liu, Q., Mura, K., & Sanchez, A. (2008). Articulatory preparation in the delayed naming task. Journal of Memory and Language, 58(2), 347-365.

Kessler, B., Treiman, R., & Mullennix, J. (2002). Phonetic Biases in Voice Key Response Time Measurements. Journal Of Memory And Language, 47(1), 145-171. http://dx.doi.org/10.1006/jmla.2001.2835

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. Behavior Research Methods, 42(3), 643-650. http://dx.doi.org/10.3758/brm.42.3.643

Kolmogorov, A. N. (1933). Sulla determinazione empirica di una legge di distribuzione. na.

Kootstra, G., van Hell, J., & Dijkstra, T. (2010). Syntactic alignment and shared word order in code-switched sentence production: Evidence from bilingual monologue and dialogue. Journal Of Memory And Language, 63(2), 210-231. http://dx.doi.org/10.1016/j.jml.2010.03.006

KOOTSTRA, G., VAN HELL, J., & DIJKSTRA, T. (2012). Priming of code-switches in sentences: The role of lexical repetition, cognates, and language proficiency. Bilingualism, 15(04), 797-819. http://dx.doi.org/10.1017/s136672891100068x

Kroll, J., & Merves, J. (1986). Lexical access for concrete and abstract words. Journal Of Experimental Psychology: Learning, Memory, And Cognition, 12(1), 92-107. http://dx.doi.org/10.1037//0278-7393.12.1.92

Kroll, J., & Stewart, E. (1994). Category Interference in Translation and Picture Naming: Evidence for Asymmetric Connections Between Bilingual Memory Representations. Journal Of Memory And Language, 33(2), 149-174. http://dx.doi.org/10.1006/jmla.1994.1008

Kroll, J., Michael, E., Tokowicz, N., & Dufour, R. (2002). The development of lexical fluency in a second language. Second Lang Res, 18(2), 137-171. http://dx.doi.org/10.1191/0267658302sr201oa

KROLL, J., VAN HELL, J., TOKOWICZ, N., & GREEN, D. (2010). The Revised Hierarchical Model: A critical review and assessment. Bilingualism, 13(03), 373-381. http://dx.doi.org/10.1017/s136672891000009x

Kucera, H., & Francis, W.

Kullback, S., & Leibler, R. (1951). On Information and Sufficiency. Ann. Math. Statist., 22(1), 79-86. http://dx.doi.org/10.1214/aoms/1177729694

LEHTONEN, M., & LAINE, M. (1999). How word frequency affects morphological processing in monolinguals and bilinguals This study was financially supported by a grant (20010) from the Joint Committee of the Nordic Social Science Research Councils (NOS-S) to the second author. Bilingualism, 6(3), 213-225. http://dx.doi.org/10.1017/s1366728903001147

Levenshtein, V. I. (1965). Binary codes with correction for deletions and insertions of the symbol 1. Problemy Peredachi Informatsii, 1(1), 12-25.

Levenshtein, V. I. (1966, February). Binary codes capable of correcting deletions, insertions, and reversals. In Soviet physics doklady (Vol. 10, No. 8, pp. 707-710).

Lopes, R., Reid, I., & Hobson, P. (2007). The two-dimensional Kolmogorov-Smirnov test. In XI International Workshop on Advanced Computing and Analysis Techniques in Physics Research. Amsterdam, Netherlands.

Lorenz, D. (2015). Predicative adjectives and the correlation of speech rate and word frequency. Presentation, ICAME 36. Trier, 31 May, 2015.

Lovatt, P., Avons, S., & Masterson, J. (2000). The Word-length Effect and Disyllabic Words. The Quarterly Journal Of Experimental Psychology Section A, 53(1), 1-22. http://dx.doi.org/10.1080/713755877

Maekawa, K., Yamazaki, M., Ogiso, T., Maruyama, T., Ogura, H., Kashino, W., ... & Den, Y. (2008, January). Balanced Corpus of Contemporary Written Japanese. In IJCNLP (pp. 101-102).

Matsumoto, A., Iidaka, T., Haneda, K., Okada, T., & Sadato, N. (2005). Linking semantic priming effect in functional MRI and event-related potentials. Neuroimage, 24(3), 624-634.

McClelland, J., & Elman, J. (1986). The TRACE model of speech perception. Cognitive Psychology, 18(1), 1-86. http://dx.doi.org/10.1016/0010-0285(86)90015-0

McClelland, J., & Rumelhart, D. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. Psychological Review, 88(5), 375-407. http://dx.doi.org/10.1037//0033-295x.88.5.375

Meuter, R., & Allport, A. (1999). Bilingual Language Switching in Naming: Asymmetrical Costs of Language Selection. Journal Of Memory And Language, 40(1), 25-40. http://dx.doi.org/10.1006/jmla.1998.2602

Meyer, D., & Schvaneveldt, R. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. Journal Of Experimental Psychology, 90(2), 227-234. http://dx.doi.org/10.1037/h0031564

Moscoso del Prado Martín, F., Kostić, A., & Baayen, R. (2004). Putting the bits together: an information theoretical perspective on morphological processing. Cognition, 94(1), 1-18. http://dx.doi.org/10.1016/j.cognition.2003.10.015

Mulder, K. (2013). Family and neighbourhood relations in the mental lexicon: A cross-language perspective (Doctoral Dissertation). Radboud Universiteit Nijmegen.

Muysken, P. (2000/2005). Bilingual speech: A typology of code-mixing (Vol. 11). Cambridge University Press.

Muysken, P. (2004). Two linguistic systems in contact: Grammar, phonology and lexicon. The handbook of bilingualism, 147-68.

Muysken, PIETER (2005). A modular Approach to Grammatical Categories: Evidence from Language Diversity and Contact. Handbook of Categorization in Cognitive Science, Edited by Henri Cohen and Claire Lefebvre _9 2005 Elsevier Ltd, 45-69.

Neath, I., Bireta, T., & Surprenant, A. (2003). The time-based word length effect and stimulus set specificity. Psychonomic Bulletin & Review, 10(2), 430-434. http://dx.doi.org/10.3758/bf03196502

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner; G. W. Humphreys (Eds.), Basic processes in reading: Visual word recognition (pp. 264-336). Hillsdale, NJ, England: Lawrence Erlbaum Associates.

Nerbonne, J., & Heeringa, W. (1997). Measuring Dialect Distance Phonetically.

New, B., ferrand, L., pallier, C., & brysbaert, M. (2006). Reexamining the word length effect in visual word recognition: New evidence from the English Lexicon Project. Psychonomic Bulletin & Review, 13(1), 45-52. http://dx.doi.org/10.3758/bf03193811

Norris, D. (2006). The Bayesian reader: Explaining word recognition as an optimal Bayesian decision process. Psychological Review, 113(2), 327-357. http://dx.doi.org/10.1037/0033-295x.113.2.327

Cognition and Lexical Access. HANDBOOK OF BILINGUALISM, 179.

Okada, K. (2013). IS OMEGA SQUARED LESS BIASED? A COMPARISON OF THREE MAJOR EFFECT SIZE INDICES IN ONE-WAY ANOVA. Behaviormetrika, 40(2), 129-147. http://dx.doi.org/10.2333/bhmk.40.129

O'regan, J. K., & Jacobs, A. M. (1992). Optimal viewing position effect in word recognition: A challenge to current theory. Journal of Experimental Psychology: Human Perception and Performance, 18(1), 185.

Palo, P., Schaeffler, S., & Scobbie, J. (2015). EFFECT OF PHONETIC ONSET ON ACOUSTIC AND ARTICULATORY SPEECH REACTION TIMES STUDIED WITH TONGUE ULTRASOUND.

Peacock, J. (1983). Two-dimensional goodness-of-fit testing in astronomy. Monthly Notices Of The Royal Astronomical Society, 202(3), 615-627. http://dx.doi.org/10.1093/mnras/202.3.615

Perea, M., & Carreiras, M. (1998). Effects of syllable frequency and syllable neighborhood frequency in visual word recognition. Journal of Experimental Psychology: Human perception and performance, 24(1), 134.

Perry, C., Ziegler, J., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. Psychological Review, 114(2), 273-315. http://dx.doi.org/10.1037/0033-295x.114.2.273

Piantadosi, S., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. Proceedings Of The National Academy Of Sciences, 108(9), 3526-3529. http://dx.doi.org/10.1073/pnas.1012551108

Pollatsek, A., Juhasz, B., Reichle, E., Machacek, D., & Rayner, K. (2008). Immediate and delayed effects of word frequency and word length on eye movements in reading: A reversed delayed effect of word length. Journal Of Experimental Psychology: Human Perception And Performance, 34(3), 726-750. http://dx.doi.org/10.1037/0096-1523.34.3.726

Pruijn, L. (2015). From ‘Daar’ to ‘There’ and Back Again: About the Mechanisms underlying Word Translation (Masters Thesis). Radboud Universiteit Nijmegen.

Pylkkänen, L., Feintuch, S., Hopkins, E., & Marantz, A. (2004). Neural correlates of the effects of morphological family frequency and family size: an MEG study. Cognition, 91(3), B35-B45. http://dx.doi.org/10.1016/j.cognition.2003.09.008

Rastle, K., Croot, K., Harrington, J., & Coltheart, M. (2005). Characterizing the Motor Execution Stage of Speech Production: Consonantal Effects on Delayed Naming Latency and Onset Duration. Journal Of Experimental Psychology: Human Perception And Performance, 31(5), 1083-1095. http://dx.doi.org/10.1037/0096-1523.31.5.1083

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. Psychological Bulletin, 124(3), 372-422. doi:10.1037/0033-2909.124.3.372

Rey, A. (1998). Orthographie et Phonologie dans la Perception des Mots ecrits. (Doctoral Dissertation). Universite de Provence Aix, Marseille.

Richardson, J. (1975). Concreteness and imageability. Quarterly Journal Of Experimental Psychology, 27(2), 235-249. http://dx.doi.org/10.1080/14640747508400483

Richardson, J. (1976). Imageability and concreteness. Bulletin Of The Psychonomic Society, 7(5), 429-431. http://dx.doi.org/10.3758/bf03337237

Richter, K. (1999). A functional units model of visual word recognition (Doctoral dissertation).

Riecker, A., Brendel, B., Ziegler, W., Erb, M., & Ackermann, H. (2008). The influence of syllable onset complexity and syllable frequency on speech motor control. Brain and language, 107(2), 102-113.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. Cognition, 42(1-3), 107-142. http://dx.doi.org/10.1016/0010-0277(92)90041-f

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. Cognition, 64(3), 249-284. http://dx.doi.org/10.1016/s0010-0277(97)00027-9

Roelofs, A. (2003). Shared phonological encoding processes and representations of languages in bilingual speakers. Language And Cognitive Processes, 18(2), 175-204. http://dx.doi.org/10.1080/01690960143000515

ROELOFS, A., DIJKSTRA, T., & GERAKAKI, S. (2012). Modeling of word translation: Activation flow from concepts to lexical items. Bilingualism, 16(02), 343-353. http://dx.doi.org/10.1017/s1366728912000612

Rumelhart, D., & McClelland, J. (1982). An interactive activation model of context effects in letter perception: II. The contextual enhancement effect and some tests and extensions of the model. Psychological Review, 89(1), 60-94. http://dx.doi.org/10.1037//0033-295x.89.1.60

Sabourin, L., Leclerc, J., Burkholder, M., & Brien, C. (2015). BILINGUAL LEXICAL ORGANIZATION: IS THERE A SENSITIVE PERIOD?. In Actes du congrès annuel de l'Association canadienne de linguistique 2014. Proceedings of the 2014 annual conference of the Canadian Linguistic Association.. Association Canadienne De Linguistique.

Scheerer, E. (1986). Orthography and lexical access. New trends in graphemics and orthography, 262-286.

Schepens, J. (2015). Bridging Linguistic Gaps: The Effects of Linguistic Distance on the Adult Learnability of Dutch as an Additional Language (Doctoral Dissertation). Radboud Universiteit Nijmegen.

SCHEPENS, J., DIJKSTRA, T., & GROOTJEN, F. (2011). Distributions of cognates in Europe as based on Levenshtein distance. Bilingualism, 15(01), 157-166. http://dx.doi.org/10.1017/s1366728910000623

Schilling, H., Rayner, K., & Chumbley, J. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. Mem Cogn, 26(6), 1270-1281. http://dx.doi.org/10.3758/bf03201199

Schreuder, R., & Baayen, R. (1997). How Complex Simplex Words Can Be. Journal Of Memory And Language, 37(1), 118-139. http://dx.doi.org/10.1006/jmla.1997.2510

Schreuder, R., & Weltens, B. (1993). The Bilingual lexicon. Amsterdam: J. Benjamins Pub. Co.

Seidenberg, M. S. (1987). Sublexical structures in visual word recognition: Access units or orthographic redundancy?.

Seidenberg, M. S. (1989). Reading complex words. In Linguistic structure in language processing (pp. 53-105). Springer Netherlands.

Sholl, A., Sankaranarayanan, A., & Kroll, J. (1995). TRANSFER BETWEEN PICTURE NAMING AND TRANSLATION:. A Test of Asymmetries in Bilingual Memory. Psychological Science, 6(1), 45-49. http://dx.doi.org/10.1111/j.1467-9280.1995.tb00303.x

Sigurd, B., Eeg-Olofsson, M., & van Weijer, J. (2004). Word length, sentence length and frequency - Zipf revisited. Studia Linguistica, 58(1), 37-52. http://dx.doi.org/10.1111/j.0039-3193.2004.00109.x

Simpson, G., Peterson, R., Casteel, M., & Burgess, C. (1989). Lexical and sentence context effects in word recognition. Journal Of Experimental Psychology: Learning, Memory, And Cognition, 15(1), 88-97. http://dx.doi.org/10.1037//0278-7393.15.1.88

Smirnov, N. (1948). Table for estimating the goodness of fit of empirical distributions. The annals of mathematical statistics, 279-281.

Swinney, D. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. Journal Of Verbal Learning And Verbal Behavior, 18(6), 645-659. http://dx.doi.org/10.1016/s0022-5371(79)90355-4

Szubko-Sitarek, W. (2015). Multilingual lexical recognition in the mental lexicon of third language users.

Team, R. C. (2015). R: A Language and environment for statistical computing [computer software]. R Foundation for Statistical Computing, Vienna, Austria.

Tham, W. W., Liow, S. J. R., Rajapakse, J. C., Leong, T. C., Ng, S. E., Lim, W. E., & Ho, L. G. (2005). Phonological processing in Chinese–English bilingual biscriptals: An fMRI study. NeuroImage, 28(3), 579-587.

Van Hell, J. G., & De Groot, A. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. Bilingualism: Language and Cognition, 1(03), 193-211.

Van Hell, J.G., & De Groot, A. M. (1998). Disentangling Context Availability and Concreteness in Lexical Decision and W ord Translation. The Quarterly Journal of Experimental Psychology: Section A, 51(1), 41-63.

Van Heuven, W. J. B., & Dijkstra, A. (2001). The semantic, orthographic, and phonological interactive activation model. In Poster presented at the 12th Conference of the European Society for Cognitive Psychology, Edinburgh, Scotland.

van Heuven, W., & Dijkstra, T. (2010). Language comprehension in the bilingual brain: fMRI and ERP support for psycholinguistic models. Brain Research Reviews, 64(1), 104-122. http://dx.doi.org/10.1016/j.brainresrev.2010.03.002

van Heuven, W., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. The Quarterly Journal Of Experimental Psychology, 67(6), 1176-1190. http://dx.doi.org/10.1080/17470218.2013.850521

Wetzel, L. (2009). Types and tokens. Cambridge, Mass.: MIT Press.

Wilson, M., & Emmorey, K. (1998). A “word length effect”for sign language: Further evidence for the role of language in structuring working memory. Mem Cogn, 26(3), 584-590. http://dx.doi.org/10.3758/bf03201164

Xiang, H. (2012). The Language networks of the brain (Doctoral Dissertation). Radboud Universiteit Nijmegen.

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart’s N: A new measure of orthographic similarity. Psychonomic Bulletin & Review, 15(5), 971-979. http://dx.doi.org/10.3758/pbr.15.5.971

Zipf, G. (1935). The psycho-biology of language. Boston: Houghton Mifflin.

Zipf, G. K. (1949). Human behavior and the principle of least effort.

11 Appendix

Appendix information is found here: section 11.1 shows Multilink's resting-level activation and Normalized Levenshtein Distance formulas; section 11.2 shows the Model-Approximate χ2 formula; section 11.3 exhibits the full dataset used for this study. Section 11.4 shows tables of descriptive statistics for each of the 11 test conditions used in the correlational tests (except the mean-regressed correlations); section 11.5 contains tables of numeric correlation output; section 11.6 contains tables of ANOVA output; section 11.7 contains tables of generalized regression output; finally, section 11.8 shows the entire R statistics script used for the results, figures, and tables.

11.1 **Multilink's Activation Functions**

i. *Resting Level Activation*

$$Rest_{MIN} + RANK * \left(\frac{|REST_{MIN} - REST_{MAX}|}{Rank_{MAX}}\right)$$

"Resting level activation = MINREST + RANK * (Abs(MINREST - MAXREST) / MAX RANK). Here MINREST is the parameter that determines the minimal resting level activation (set at -.05). RANK is the ranking of the word in the frequency ordered item list for the language in question. Abs() computes the absolute value. MAXREST is the parameter that determines the maximal resting level activation (set at .0). Finally, MAXRANK is the highest rank that occurs in the frequency ordered item list for the language. Words in the same language and of the same frequency have the same RANK; as a consequence, they also have the same resting level activation."

(Dijkstra et al., in prep.: 27)

ii. *Normalized Levenshtein Distance*

$$1 - (Distance / Length)$$

Where *Distance* is defined as:

$$\wedge LD$$

And where *Length* is defined as:

$$\vee LENGTH$$

"When a word (of any length) is presented to Multilink for simulation, it leads to the activation of lexical-orthographic representations from several languages depending on their (1) orthographic similarity with the input, and (2) subjective frequency of usage. The orthographic similarity of the input word to stored lexical representations is computed by determining the Levenshtein distance between the involved letter strings (see Schepens, Dijkstra, & Grootjen, 2011). This metric involves the computation of the minimal number of deletions, substitutions, and insertions needed to edit one expression into the other. The Levenshtein distance is in use by researchers with different backgrounds to explore the relation between orthographic, phonetic, and cross-linguistic similarity (e.g., Heeringa, 2004; Kessler, 2005; Levenshtein, 1966). By normalizing the Levenshtein distance for word length, the activation of word candidates of different lengths is possible. The formula to compute the normalized Levenshtein score for two expressions is the following: score = 1 - (distance / length), where length = max(length of source expression, length of destination expression) and distance = min(number of insertions, deletions, and substitutions)."

(Djikstra et al., in prep.: 30)

11.2 Model-Approximate χ2 Test Formula

$$\chi^2 = \sum_{i}^{N} \frac{(RT_{DATA_i} - RT_{MODEL_i})^2}{\left(\frac{S^2_{DATA_i}}{n_i}\right) + \left(\frac{S^2_{MODEL_i}}{m_i}\right)}$$

"A second common stress-measure is the sum of the squared differences between the predicted and obtained RTs. This measure can be turned into an approximated x2 statistic by taking into account the sample variance s2 in the obtained and predicted RTs as follows [. . .] [*Formula here*] [. . .] When the model values are based on the use of p free parameters, $Y^M_{i\,j}$ are no longer independent for all i,j. The number of degrees of freedom for the x2-distribution then becomes N-p, instead of N (where N indicates the number of data points involved). Both small and large simulations were run. For the small simulations, sample size m; was set equal to n; (leading to a simplification of the formula); for the large simulations sample size m; was set at

35,000. For the computation of means and variances, data of all subjects were pooled. This has some consequences. As can be observed on the basis of Equation (3), considering subjects as replications on the one hand leads to an over-estimation of the population variance (due to subject effects), and thus results in a denominator that is too large. In fact, the variance in the pooled data was much larger than in the model. On the other hand, pooling leads to a much larger n;, which results in a smaller denominator. The first consideration alone would lead to a x2 that is smaller than it should be and to too few rejections of the model; the second leads to too many rejections. A better, but more time consuming, approach would be to fit the model to each individual subject, compute the x2 for each subject, and sum the x2s over subjects, while adapting the degrees of freedom."

(Dijkstra, 1990: 163-165)

11.3 **Data Set**

This section displays the final dataset, as written at the end of statistical testing (see section 11.9 to view the statistical script). Bolded column titles were computed in R for each item, and were not part of the original data file, derived from Pruijn (2015). Incorrect and inaccurate stimulus are not included.

	Stimuli	Translation	Latency	cycles	**P.Latency.lm**	**P.Latency.zm**	**Stim.length**	freq	**Levdist**	concreteness	Mean.N.RT	category	Direction	cognacy	freq.cat	Ph.Onset
1	angry	boos	772	27.01	862	903	5	65	5	2.53	495.82	Backward HFNcog	Backward	Noncognate	High-FREQ	voiced
2	ant	mier	1099	30.77	951	1166	3	4	4	4.86	538.27	Backward LFNcog	Backward	Noncognate	Low-FREQ	voiced
3	art	kunst	935	26.25	844	857	3	211	4	4.17	507.76	Backward HFNcog	Backward	Noncognate	High-FREQ	voiceless
4	autumn	herfst	1061	27.25	867	917	6	34	6	3.27	558.76	Backward LFNcog	Backward	Noncognate	Low-FREQ	voiceless
5	axe	bijl	968	27.44	872	929	3	7	4	5	519.05	Backward LFNcog	Backward	Noncognate	Low-FREQ	voiced
6	baker	bakker	900	22.60	757	668	5	36	1	4.71	568.42	Backward LFCog	Backward	Cognate	Low-FREQ	voiced
7	bible	bijbel	774	27.12	864	909	5	61	3	4.61	599.88	Backward LFCog	Backward	Cognate	Low-FREQ	voiced
8	black	zwart	787	26.24	843	856	5	224	4	3.76	536.58	Backward HFNcog	Backward	Noncognate	High-FREQ	voiced
9	boat	boot	737	22.89	764	681	4	56	1	4.93	545.18	Backward HFCog	Backward	Cognate	High-FREQ	voiced
10	body	lichaam	872	25.43	824	810	4	283	7	4.79	580.58	Backward HFNcog	Backward	Noncognate	High-FREQ	voiced
11	book	boek	660	21.76	737	631	4	269	1	4.9	521.36	Backward HFCog	Backward	Cognate	High-FREQ	voiced
12	boss	baas	781	27.11	864	909	4	22	2	3.83	525.6	Backward	Back	Cognat	High-	voice

											8	HFCog	ward	e	FREQ	d
Stimuli	Translation	Latency	cycles	**P.Latency.lm**	**P.Latency.zm**	**Stim.length**	freq	**Lev dist**	concreteness	Mean.N.RT	category	Direction	cognacy	freq.cat	Ph.Onset	Stimuli
13	boy	jongen	828	25.84	834	833	3	207	5	4.76	549.63	Backward HFNcog	Backward	Noncognate	High-FREQ	voiced
14	bride	bruid	910	24.15	794	743	5	34	2	4.63	583.18	Backward LFCog	Backward	Cognate	Low-FREQ	voiced
15	brown	bruin	739	23.63	781	717	5	183	2	4.48	570.53	Backward LFCog	Backward	Cognate	Low-FREQ	voiced
16	building	gebouw	809	26.56	851	875	8	170	8	4.64	597.11	Backward HFNcog	Backward	Noncognate	High-FREQ	voiceless
17	bull	stier	1100	27.37	870	925	4	19	5	4.85	533.03	Backward LFNcog	Backward	Noncognate	Low-FREQ	voiceless
18	car	auto	765	25.62	829	821	3	276	4	4.89	507.56	Backward HFNcog	Backward	Noncognate	High-FREQ	vowel
19	clean	schoon	821	26.96	861	900	5	72	4	3.07	489.33	Backward HFNcog	Backward	Noncognate	High-FREQ	voiceless
20	cloud	wolk	874	27.30	869	921	5	30	4	4.54	559.15	Backward LFNcog	Backward	Noncognate	Low-FREQ	voiced
21	coffee	koffie	710	23.11	769	692	6	80	2	4.81	534.59	Backward HFCog	Backward	Cognate	High-FREQ	voiceless
22	coral	koraal	934	23.78	785	724	5	2	2	4.4	532.74	Backward LFCog	Backward	Cognate	Low-FREQ	voiceless
23	coward	lafaard	1138	27.47	873	931	6	5	4	2.93	567.35	Backward LFNcog	Backward	Noncognate	Low-FREQ	voiced
24	culture	cultuur	737	22.94	765	684	7	59	2	2.04	526.18	Backward LFCog	Backward	Cognate	Low-FREQ	voiceless
25	curse	vloek	1024	23.95	789	733	5	11	5	2.39	515.36	Backward LFNcog	Backward	Noncognate	Low-FREQ	voiceless
26	dark	donker	787	27.06	863	906	4	37	4	4.29	551.21	Backward HFNcog	Backward	Noncognate	High-FREQ	voiced

Stimuli	Translation	Latency	cycles	**P.Latency.lm**	**P.Latency.zm**	**Stim.length**	freq	**Lev dist**	concreteness	Mean.N.RT	category	Direction	cognacy	freq.cat	Ph.Onset	voiced
28	dirty	vies	919	27.27	868	919	5	37	4	4.23	515.05	Backward HFNcog	Backward	Noncognate	High-FREQ	voiced
29	dog	hond	744	26.70	854	884	3	77	3	4.85	525.18	Backward HFNcog	Backward	Noncognate	High-FREQ	voiceless
30	domain	domein	901	22.66	758	671	6	9	1	3.4	560.54	Backward LFCog	Backward	Cognate	Low-FREQ	voiced
31	ear	oor	732	27.18	866	913	3	30	2	5	520.9	Backward LFCog	Backward	Cognate	Low-FREQ	vowel
32	eye	oog	669	26.15	841	851	3	143	3	4.9	476.03	Backward HFNcog	Backward	Noncognate	High-FREQ	vowel
33	face	gezicht	770	24.97	813	785	4	386	6	4.87	516.95	Backward HFNcog	Backward	Noncognate	High-FREQ	voiceless
34	farmer	boer	965	30.45	944	1141	6	31	4	4.54	584.49	Backward LFNcog	Backward	Noncognate	Low-FREQ	voiced
35	father	vader	764	25.28	821	802	6	188	3	4.52	538.13	Backward HFCog	Backward	Cognate	High-FREQ	voiced
36	feeling	gevoel	846	27.79	880	952	7	61	6	1.68	542.37	Backward HFNcog	Backward	Noncognate	High-FREQ	voiceless
37	fever	koorts	1061	27.29	868	920	5	26	6	3.27	553.3	Backward LFNcog	Backward	Noncognate	Low-FREQ	voiceless
38	field	veld	877	23.70	783	720	5	127	2	4.26	529.39	Backward HFCog	Backward	Cognate	High-FREQ	voiced
39	figure	figuur	807	23.00	766	687	6	210	2	3.63	530.58	Backward HFCog	Backward	Cognate	High-FREQ	voiceless
40	flame	vlam	849	24.19	794	745	5	19	2	4.67	544.84	Backward LFCog	Backward	Cognate	Low-FREQ	voiced
41	foot	voet	767	26.60	852	878	4	113	2	4.9	492.03	Backward HFCog	Backward	Cognate	High-FREQ	voiced
42	fox	vos	889	27.42	872	928	3	16	2	4.97	553.3	Backward	Back	Cognat	Low-	voice

												LFCog	ward	e	FREQ	d
Stimuli	Translation	Latency	cycles	**P.Latency.lm**	**P.Latency.zm**	**Stim.length**	freq	**Lev dist**	concreteness	Mean.N.RT	category	Direction	cognacy	freq.cat	Ph.Onset	voiced
44	garden	tuin	830	26.30	845	860	6	164	5	4.73	519.33	Backward LFNcog	Backward	Noncognate	Low-FREQ	voiceless
45	garlic	knoflook	1027	27.44	872	929	6	4	7	4.89	577.37	Backward LFNcog	Backward	Noncognate	Low-FREQ	voiceless
46	girl	meisje	794	25.65	829	823	4	225	5	4.85	542.03	Backward HFNcog	Backward	Noncognate	High-FREQ	voiced
47	glove	handschoen	1102	27.47	873	931	5	5	9	4.97	551.22	Backward LFNcog	Backward	Noncognate	Low-FREQ	voiceless
48	ground	grond	737	21.32	726	612	6	193	1	4.77	481.51	Backward HFCog	Backward	Cognate	High-FREQ	voiceless
49	hair	haar	670	22.61	757	668	4	191	1	4.97	510.79	Backward HFCog	Backward	Cognate	High-FREQ	voiceless
50	healthy	gezond	908	27.16	865	912	7	33	6	3.31	532	Backward LFNcog	Backward	Noncognate	Low-FREQ	voiceless
51	honey	honing	974	27.35	870	924	5	21	3	4.88	538.41	Backward HFCog	Backward	Cognate	High-FREQ	voiceless
52	horse	paard	907	26.71	855	884	5	85	5	5	510.55	Backward HFNcog	Backward	Noncognate	High-FREQ	voiceless
53	house	huis	701	24.78	809	775	5	479	3	5	510.28	Backward HFCog	Backward	Cognate	High-FREQ	voiceless
54	hunter	jager	976	27.41	871	928	6	11	4	4.41	543.75	Backward LFNcog	Backward	Noncognate	Low-FREQ	voiced
55	juice	sap	997	27.37	870	925	5	20	5	4.89	546.79	Backward LFNcog	Backward	Noncognate	Low-FREQ	voiceless
56	key	sleutel	818	26.94	860	898	3	92	6	4.89	483.22	Backward HFNcog	Backward	Noncognate	High-FREQ	voiceless
57	kidney	nier	925	27.47	873	931	6	5	4	4.96	523.08	Backward LFNcog	Backward	Noncognate	Low-FREQ	voiced

Sti mul i	Transl ation	Latenc y	cycl es	**P.Laten cy.lm**	**P.Laten cy.zm**	**Stim.le ngth**	freq	**Lev dist**	concret eness	Mean.N .RT	catego ry	Direction	cogn acy	freq.ca t	Ph.On set	voice d
59	metho d	metho de	712	22.12	745	647	6	77	1	2.41	571.9 3	Backward LFCog	Back ward	Cognat e	Low-FREQ	voice d
60	middl e	midde n	871	22.76	760	676	6	171	2	3.69	507.3 4	Backward HFCog	Back ward	Cognat e	High-FREQ	voice d
61	minut e	Minuu t	668	23.47	777	709	6	74	3	3.04	508.4 1	Backward HFCog	Back ward	Cognat e	High-FREQ	voice d
62	mirror	spiegel	931	27.11	864	909	6	41	6	4.97	531.1 1	Backward LFNcog	Back ward	Nonco gnate	Low-FREQ	voice less
63	mone y	geld	810	25.21	819	798	5	403	5	4.54	520.9 5	Backward HFNcog	Back ward	Nonco gnate	High-FREQ	voice less
64	morni ng	ochten d	971	26.36	846	864	7	214	6	3.44	563.1 5	Backward HFNcog	Back ward	Nonco gnate	High-FREQ	vowe l
65	mothe r	moede r	654	21.92	740	638	6	221	2	4.6	506.7 2	Backward HFCog	Back ward	Cognat e	High-FREQ	voice d
66	motiv e	motief	886	23.69	783	720	6	14	2	1.5	575.7 9	Backward LFCog	Back ward	Cognat e	Low-FREQ	voice d
67	mouse	muis	730	27.42	872	928	5	10	3	4.83	560.7 2	Backward LFCog	Back ward	Cognat e	Low-FREQ	voice d
68	mouth	mond	726	25.99	837	842	5	134	3	4.74	499.8 9	Backward HFCog	Back ward	Cognat e	High-FREQ	voice d
69	muscl e	spier	946	27.28	868	919	6	44	5	4.5	587.3 2	Backward LFNcog	Back ward	Nonco gnate	Low-FREQ	voice less
70	music	Muziek	681	26.09	840	848	5	221	4	4.31	534.3 7	Backward HFCog	Back ward	Cognat e	High-FREQ	voice d
71	name	naam	631	25.32	822	805	4	306	2	3.5	530.9 5	Backward HFCog	Back ward	Cognat e	High-FREQ	voice d
72	night	nacht	709	22.47	753	662	5	428	2	4.52	508.8 4	Backward HFCog	Back ward	Cognat e	High-FREQ	voice d
73	office	kantoo	103	26.28	844	859	6	249	7	4.93	537.9	Backward	Back	Nonco	High-	voice

		r	9								5	HFNcog	ward	gnate	FREQ	less
Sti mul i	Transl ation	Latenc y	cycl es	**P.Laten cy.lm**	**P.Laten cy.zm**	**Stim.le ngth**	freq	**Lev dist**	concret eness	Mean.N .RT	catego ry	Direction	cogn acy	freq.ca t	Ph.On set	voice less
75	panic	paniek	724	23.61	781	716	5	26	2	3.04	518.8 7	Backward LFCog	Back ward	Cognat e	Low- FREQ	voice less
76	peach	perzik	937	27.49	873	933	5	3	4	4.9	552.2 5	Backward LFNcog	Back ward	Nonco gnate	Low- FREQ	voice less
77	pearl	parel	869	24.26	796	748	5	12	2	4.87	559.3 5	Backward LFCog	Back ward	Cognat e	Low- FREQ	voice less
78	pencil	potloo d	973	27.39	871	926	6	15	6	4.88	509.5 5	Backward LFNcog	Back ward	Nonco gnate	Low- FREQ	voice less
79	peppe r	peper	776	22.66	758	671	6	7	1	4.59	617.3	Backward LFCog	Back ward	Cognat e	Low- FREQ	voice less
80	pipe	pijp	796	27.32	869	922	4	20	2	4.88	552.7 5	Backward LFCog	Back ward	Cognat e	Low- FREQ	voice less
81	pirate	piraat	711	23.78	785	724	6	3	2	4.64	554.7	Backward LFCog	Back ward	Cognat e	Low- FREQ	voice less
82	plate	bord	120 3	27.21	867	915	5	24	5	4.77	526.7 8	Backward LFNcog	Back ward	Nonco gnate	Low- FREQ	voice d
83	police	politie	729	22.59	756	668	6	206	2	4.79	510.7 3	Backward HFCog	Back ward	Cognat e	High- FREQ	voice less
84	questi on	vraag	812	25.15	817	795	8	259	8	3.36	545.5 5	Backward HFNcog	Back ward	Nonco gnate	High- FREQ	voice d
85	rabbit	konijn	873	27.41	871	928	6	11	6	4.93	540.2 4	Backward LFNcog	Back ward	Nonco gnate	Low- FREQ	voice less
86	raw	rauw	891	23.03	767	688	3	43	1	3.35	546.5 8	Backward LFCog	Back ward	Cognat e	Low- FREQ	voice d
87	rice	rijst	856	27.35	870	924	4	27	3	4.86	523.2 9	Backward LFCog	Back ward	Cognat e	Low- FREQ	voice d
88	rich	rijk	722	26.76	856	887	4	79	2	2.79	557.5	Backward HFCog	Back ward	Cognat e	High- FREQ	voice d

Stimuli	Translation	Latency	cycles	**P.Latency.lm**	**P.Latency.zm**	**Stim.length**	freq	**Lev dist**	concreteness	Mean.N.RT	category	Direction	cognacy	freq.cat	Ph.Onset	voiceless
90	scar	litteken	1030	27.44	872	929	4	7	8	4.74	560.5	Backward LFNcog	Backward	Noncognate	Low-FREQ	voiced
91	science	wetenschap	1003	26.67	854	882	7	91	7	2.79	604.66	Backward LFNcog	Backward	Noncognate	Low-FREQ	voiced
92	shark	haai	916	27.44	872	929	5	14	3	4.93	563.91	Backward LFNcog	Backward	Noncognate	Low-FREQ	voiceless
93	shower	douche	842	28.41	895	993	6	15	6	4.89	538	Backward HFNcog	Backward	Noncognate	High-FREQ	voiced
94	silver	zilver	787	22.61	757	669	6	31	1	4.52	527.61	Backward LFCog	Backward	Cognate	Low-FREQ	voiced
95	slave	slaaf	839	24.18	794	744	5	16	2	4.38	496.82	Backward LFCog	Backward	Cognate	Low-FREQ	voiceless
96	sleeve	mouw	1140	27.39	871	927	6	11	6	4.84	591.54	Backward LFNcog	Backward	Noncognate	Low-FREQ	voiced
97	soft	zacht	838	26.70	854	884	4	69	4	3.88	515.71	Backward LFNcog	Backward	Noncognate	Low-FREQ	voiced
98	song	liedje	965	28.31	893	986	4	33	6	4.46	496.18	Backward HFNcog	Backward	Noncognate	High-FREQ	voiced
99	space	ruimte	921	26.12	841	849	5	185	5	3.54	521.63	Backward HFNcog	Backward	Noncognate	High-FREQ	voiced
100	spicy	pittig	1072	27.48	873	932	5	2	5	3.31	562.48	Backward LFNcog	Backward	Noncognate	Low-FREQ	voiceless
101	spoon	lepel	838	28.45	896	996	5	11	5	4.96	535.18	Backward LFNcog	Backward	Noncognate	Low-FREQ	voiced
102	story	verhaal	886	26.09	840	848	5	167	7	3.3	557.03	Backward HFNcog	Backward	Noncognate	High-FREQ	voiced
103	street	straat	752	22.60	757	668	6	253	2	4.75	541.78	Backward HFCog	Backward	Cognate	High-FREQ	voiceless
104	sun	zon	714	26.78	856	889	3	123	2	4.83	509.4	Backward	Back	Cognat	High-	voice

											2	HFCog	ward	e	FREQ	d
Stimuli	Translation	Latency	cycles	**P.Latency.lm**	**P.Latency.zm**	**Stim.length**	freq	**Lev dist**	concreteness	Mean.N.RT	category	Direction	cognacy	freq.cat	Ph.Onset	voiced
106	table	tafel	707	25.89	835	836	5	204	3	4.9	573.29	Backward HFCog	Backward	Cognate	High-FREQ	voiceless
107	task	taak	860	22.60	757	668	4	65	1	2.84	539.82	Backward LFCog	Backward	Cognate	Low-FREQ	voiceless
108	tea	thee	735	27.11	864	909	3	32	2	4.69	490.28	Backward HFCog	Backward	Cognate	High-FREQ	voiceless
109	thief	dief	789	24.25	796	748	5	8	2	4.37	544.7	Backward LFCog	Backward	Cognate	Low-FREQ	voiced
110	tiger	tijger	723	22.70	759	673	5	4	1	5	543.49	Backward LFCog	Backward	Cognate	Low-FREQ	voiceless
111	tired	moe	814	27.00	862	902	5	48	4	3	579.21	Backward HFNcog	Backward	Noncognate	High-FREQ	voiced
112	trophy	trofee	1034	27.49	873	933	6	2	3	4.89	579.83	Backward LFCog	Backward	Cognate	Low-FREQ	voiceless
113	uncle	oom	819	26.77	856	888	5	59	5	4.24	558.24	Backward HFNcog	Backward	Noncognate	High-FREQ	vowel
114	vague	vaag	894	27.19	866	914	5	24	3	1.55	580.55	Backward LFCog	Backward	Cognate	Low-FREQ	voiced
115	vision	visie	987	23.46	777	709	6	58	2	3.39	537.58	Backward LFCog	Backward	Cognate	Low-FREQ	voiced
116	voice	stem	962	25.56	827	818	5	232	5	4.13	510.32	Backward HFNcog	Backward	Noncognate	High-FREQ	voiceless
117	wall	mum	808	26.39	847	866	4	173	4	4.86	532.61	Backward HFNcog	Backward	Noncognate	High-FREQ	voiced
118	warmth	warmte	788	22.48	754	663	6	28	1	3.39	602.65	Backward LFCog	Backward	Cognate	Low-FREQ	voiced
119	window	raam	992	26.45	848	869	6	132	6	4.86	514.62	Backward HFNcog	Backward	Noncognate	High-FREQ	voiced

Stimuli	Translation	Latency	cycles	**P.Latency.lm**	**P.Latency.zm**	**Stim.length**	freq	**Lev dist**	concreteness	Mean.N.RT	category	Direction	cognacy	freq.cat	Ph.Onset	voiced
121	word	woord	815	21.42	728	617	4	279	1	3.56	531.41	BackwardHFCog	Backward	Cognate	High-FREQ	voiced
122	work	werk	695	21.18	723	606	4	772	1	3.48	512.21	BackwardHFCog	Backward	Cognate	High-FREQ	voiced
123	world	wereld	730	21.67	734	627	5	816	2	4.36	530.05	BackwardHFCog	Backward	Cognate	High-FREQ	voiced
124	year	jaar	675	24.69	807	771	4	508	2	3.25	560	BackwardHFCog	Backward	Cognate	High-FREQ	voiced
125	yellow	geel	769	27.10	864	908	6	65	4	4.3	486.16	BackwardLFNcog	Backward	Noncognate	Low-FREQ	voiceless
126	auto	car	734	25.61	828	820	4	208	4	5	517.97	ForwardHFNcog	Forward	Noncognate	High-FREQ	voiceless
127	avond	evening	854	26.11	840	849	5	194	5	3	521.27	ForwardHFNcog	Forward	Noncognate	High-FREQ	vowel
128	bakker	baker	928	22.56	756	666	6	11	1	4.6	564.03	ForwardLFCog	Forward	Cognate	Low-FREQ	voiced
129	balkon	balcony	832	23.42	776	706	6	13	2	4.6	545.13	ForwardLFCog	Forward	Cognate	Low-FREQ	voiced
130	bedrijf	company	1003	26.22	843	855	7	119	7	3.27	542.16	ForwardHFNcog	Forward	Noncognate	High-FREQ	voiceless
131	bier	beer	791	22.93	765	684	4	56	1	4.73	535.78	ForwardHFCog	Forward	Cognate	High-FREQ	voiced
132	bijbel	bible	768	27.10	864	908	6	24	3	3.87	598.73	ForwardLFCog	Forward	Cognate	Low-FREQ	voiced
133	bijl	axe	1076	28.48	897	998	4	8	4	4.87	542.39	ForwardLFNcog	Forward	Noncognate	Low-FREQ	vowel
134	boek	book	779	21.74	736	630	4	250	1	4.93	524.14	ForwardHFCog	Forward	Cognate	High-FREQ	voiced
135	boos	angry	772	27.00	862	902	4	41	5	2.73	544.5	ForwardH	Forw	Nonco	High-	vowe

											3	FNcog	ard	gnate	FREQ	l
Stimuli	Translation	Latency	cycles	**P.Latency.lm**	**P.Latency.zm**	**Stim.length**	freq	**Lev dist**	concreteness	Mean.N.RT	category	Direction	cognacy	freq.cat	Ph.Onset	voiced
137	bruin	brown	712	23.39	775	705	5	40	2	3.8	531.54	ForwardLFCog	Forward	Cognate	Low-FREQ	voiced
138	bureau	desk	1010	26.86	858	893	6	82	5	4.67	550.97	ForwardHFNcog	Forward	Noncognate	High-FREQ	voiced
139	dak	roof	894	27.01	862	903	3	43	4	4.8	545.38	ForwardHFNcog	Forward	Noncognate	High-FREQ	voiced
140	dief	thief	811	24.25	796	748	4	7	2	4	535.21	ForwardLFCog	Forward	Cognate	Low-FREQ	voiceless
141	dochter	daughter	696	23.43	776	707	7	93	3	3.6	552.24	ForwardHFCog	Forward	Cognate	High-FREQ	voiced
142	domein	domain	874	22.67	758	671	6	12	1	3	632.25	ForwardLFCog	Forward	Cognate	Low-FREQ	voiced
143	donker	dark	714	30.37	942	1135	6	57	4	4	580.97	ForwardHFNcog	Forward	Noncognate	High-FREQ	voiced
144	dorp	village	886	26.76	856	888	4	98	7	4.33	558.76	ForwardHFNcog	Forward	Noncognate	High-FREQ	voiced
145	droog	dry	745	26.96	861	899	5	45	3	3.8	524.47	ForwardLFCog	Forward	Cognate	Low-FREQ	voiced
146	duif	pigeon	988	27.46	873	931	4	8	6	4.93	523.68	ForwardLFNcog	Forward	Noncognate	Low-FREQ	voiceless
147	einde	end	773	22.61	757	669	5	155	2	2.53	509.72	ForwardHFCog	Forward	Cognate	High-FREQ	vowel
148	gebouw	building	732	26.51	850	873	6	55	8	4.53	572.16	ForwardHFNcog	Forward	Noncognate	High-FREQ	voiced
149	geel	yellow	663	27.08	864	907	4	23	4	3.8	527.33	ForwardLFNcog	Forward	Noncognate	Low-FREQ	voiced
150	geit	goat	802	27.43	872	929	4	5	2	4.87	584.35	ForwardLFCog	Forward	Cognate	Low-FREQ	voiced

Stimuli	Translation	Latency	cycles	**P.Latency.lm**	**P.Latency.zm**	**Stim.length**	freq	**Lev dist**	concreteness	Mean.N.RT	category	Direction	cognacy	freq.cat	Ph.Onset	voiced
152	gevaar	danger	768	26.77	856	888	6	98	5	2.73	533.56	ForwardHFNcog	Forward	Noncognate	High-FREQ	voiced
153	gevoel	feeling	834	26.23	843	856	6	255	6	2.33	540.08	ForwardHFNcog	Forward	Noncognate	High-FREQ	voiceless
154	gezicht	face	773	29.91	931	1100	7	447	6	4.2	541.57	ForwardHFNcog	Forward	Noncognate	High-FREQ	voiceless
155	gezond	healthy	743	27.16	865	912	6	41	6	2.13	547.36	ForwardLFNcog	Forward	Noncognate	Low-FREQ	voiceless
156	goud	gold	744	22.78	761	677	4	36	1	4.73	574.29	ForwardHFCog	Forward	Cognate	High-FREQ	voiced
157	grond	ground	688	21.43	729	617	5	321	1	4.6	520.79	ForwardHFCog	Forward	Cognate	High-FREQ	voiced
158	haai	shark	822	27.43	872	929	4	1	3	4.87	561	ForwardLFNcog	Forward	Noncognate	Low-FREQ	voiceless
159	herfst	autumn	1037	27.24	867	917	6	22	6	2.73	565.03	ForwardLFNcog	Forward	Noncognate	Low-FREQ	vowel
160	hond	dog	737	26.71	855	884	4	107	3	5	531.19	ForwardHFNcog	Forward	Noncognate	High-FREQ	voiced
161	honing	honey	718	27.35	870	924	6	12	3	4.67	578.1	ForwardLFCog	Forward	Cognate	Low-FREQ	voiceless
162	hoofd	head	766	24.86	810	779	5	515	3	4.8	536.26	ForwardHFCog	Forward	Cognate	High-FREQ	voiceless
163	huis	house	714	24.79	809	776	4	541	3	4.93	536.42	ForwardHFCog	Forward	Cognate	High-FREQ	voiceless
164	ijzer	iron	930	27.09	864	908	5	17	4	4.53	568.53	ForwardLFNcog	Forward	Noncognate	Low-FREQ	vowel
165	jager	hunter	892	27.41	871	928	5	9	4	4.07	600.24	ForwardLFNcog	Forward	Noncognate	Low-FREQ	voiceless
166	jonge	boy	755	25.84	834	833	6	202	5	4.67	526.3	ForwardH	Forw	Nonco	High-	voice

	n										9	FNcog	ard	gnate	FREQ	d
Stimuli	Translation	Latency	cycles	**P.Latency.lm**	**P.Latency.zm**	**Stim.length**	freq	**Lev dist**	concreteness	Mean.N.RT	category	Direction	cognacy	freq.cat	Ph.Onset	vowel
168	kat	cat	696	23.50	778	711	3	49	1	4.87	548.9	ForwardHFCog	Forward	Cognate	High-FREQ	voiceless
169	klimaat	climate	820	24.33	798	752	7	28	3	2.27	524.74	ForwardLFCog	Forward	Cognate	Low-FREQ	voiceless
170	knop	button	1038	27.36	870	924	4	17	5	4.53	532.87	ForwardLFNcog	Forward	Noncognate	Low-FREQ	voiced
171	koffie	coffee	684	23.17	770	695	6	110	2	4.53	555.26	ForwardHFCog	Forward	Cognate	High-FREQ	voiceless
172	koorts	fever	897	27.29	868	920	6	21	6	3.87	586.81	ForwardLFNcog	Forward	Noncognate	Low-FREQ	voiceless
173	lawaai	noise	980	27.18	866	913	6	31	6	3.87	569.21	ForwardLFNcog	Forward	Noncognate	Low-FREQ	voiced
174	lepel	spoon	834	27.40	871	927	5	11	5	4.93	548.19	ForwardLFNcog	Forward	Noncognate	Low-FREQ	voiceless
175	lichaam	body	752	25.42	824	810	7	263	7	4.47	550.82	ForwardHFNcog	Forward	Noncognate	High-FREQ	voiced
176	lied	song	869	27.26	868	918	4	20	4	4.47	574.95	ForwardLFNcog	Forward	Noncognate	Low-FREQ	voiceless
177	maag	stomach	1025	27.15	865	911	4	38	5	4.53	514.76	ForwardLFNcog	Forward	Noncognate	Low-FREQ	voiceless
178	maan	moon	667	26.92	860	897	4	62	2	4.93	564.21	ForwardHFCog	Forward	Cognate	High-FREQ	voiced
179	meisje	girl	720	25.65	829	823	6	236	5	4.47	540.82	ForwardLFNcog	Forward	Noncognate	Low-FREQ	voiced
180	mes	knife	738	26.97	861	900	3	32	5	4.73	548.32	ForwardHFNcog	Forward	Noncognate	High-FREQ	voiced
181	mier	ant	1033	27.48	873	932	4	1	4	4.87	549.76	ForwardLFNcog	Forward	Noncognate	Low-FREQ	vowel

Stimuli	Translation	Latency	cycles	**P.Latency.lm**	**P.Latency.zm**	**Stim.length**	freq	**Lev dist**	concreteness	Mean.N.RT	category	Direction	cognacy	freq.cat	Ph.Onset	voiceless
183	moeder	mother	702	22.08	744	645	6	554	2	3	502.44	ForwardHFCog	Forward	Cognate	High-FREQ	voiced
184	mond	mouth	676	26.01	838	843	4	220	3	4.73	564.5	ForwardHFCog	Forward	Cognate	High-FREQ	voiced
185	mouw	sleeve	1012	27.40	871	927	4	13	6	4.8	593.74	ForwardLFNcog	Forward	Noncognate	Low-FREQ	voiceless
186	muis	mouse	688	27.42	872	928	4	9	3	5	563.87	ForwardLFCog	Forward	Cognate	Low-FREQ	voiced
187	muur	wall	816	26.37	847	864	4	90	4	4.8	531.18	ForwardHFNcog	Forward	Noncognate	High-FREQ	voiced
188	muziek	music	642	26.06	839	846	6	115	3	4	530.97	ForwardHFCog	Forward	Cognate	High-FREQ	voiced
189	naam	name	621	25.32	822	804	4	293	2	2.6	549.79	ForwardHFCog	Forward	Cognate	High-FREQ	voiced
190	nacht	night	666	22.38	751	658	5	266	2	3.8	537.54	ForwardHFCog	Forward	Cognate	High-FREQ	voiced
191	natuur	nature	636	22.78	761	676	6	91	2	2.93	525.82	ForwardLFCog	Forward	Cognate	Low-FREQ	voiced
192	nier	kidney	891	29.62	924	1078	4	3	4	4.4	567.37	ForwardLFNcog	Forward	Noncognate	Low-FREQ	voiceless
193	oor	ear	738	27.18	866	913	3	40	2	4.8	494.42	ForwardLFCog	Forward	Cognate	Low-FREQ	vowel
194	oorlog	war	726	25.35	822	806	6	184	5	2.87	522.45	ForwardHFNcog	Forward	Noncognate	High-FREQ	voiced
195	paard	horse	776	26.71	855	884	5	99	5	4.93	519.45	ForwardHFNcog	Forward	Noncognate	High-FREQ	voiceless
196	parel	pearl	863	26.05	839	846	5	2	2	4.33	599.82	ForwardLFCog	Forward	Cognate	Low-FREQ	voiceless
197	peper	pepper	775	22.67	758	671	5	13	1	4.47	584.7	ForwardLF	Forw	Cognat	Low-	voice

											8	Cog	ard	e	FREQ	less
Stimuli	Translation	Latency	cycles	**P.Latency.lm**	**P.Latency.zm**	**Stim.length**	freq	**Lev dist**	concreteness	Mean.N.RT	category	Direction	cognacy	freq.cat	Ph.Onset	voiceless
199	pijn	pain	726	26.48	849	871	4	148	2	3.29	505.59	ForwardHFCog	Forward	Cognate	High-FREQ	voiceless
200	pijp	pipe	848	27.32	869	922	4	19	2	4.53	564.47	ForwardLFCog	Forward	Cognate	Low-FREQ	voiceless
201	piraat	pirate	710	23.78	785	724	6	2	2	4.2	596.05	ForwardLFCog	Forward	Cognate	Low-FREQ	voiceless
202	pittig	spicy	1048	27.48	873	932	6	3	5	2.73	573.13	ForwardLFNcog	Forward	Noncognate	Low-FREQ	voiceless
203	prijs	price	833	23.66	782	718	5	76	2	3.21	523.77	ForwardHFCog	Forward	Cognate	High-FREQ	voiceless
204	rauw	raw	742	22.99	766	686	4	21	1	3.8	560.19	ForwardLFCog	Forward	Cognate	Low-FREQ	voiced
205	regen	rain	681	26.93	860	898	5	53	3	4.6	575.1	ForwardLFCog	Forward	Cognate	Low-FREQ	voiced
206	rijk	rich	699	26.77	856	888	4	92	2	2.13	528.92	ForwardHFCog	Forward	Cognate	High-FREQ	voiced
207	rivier	river	657	22.13	745	647	6	69	1	4.73	527.97	ForwardHFCog	Forward	Cognate	High-FREQ	voiced
208	ruimte	space	814	26.10	840	848	6	138	5	2.93	573.08	ForwardHFNcog	Forward	Noncognate	High-FREQ	voiceless
209	saus	sauce	757	24.14	793	742	4	19	2	4.4	545.18	ForwardLFCog	Forward	Cognate	Low-FREQ	voiceless
210	schoon	clean	789	26.95	860	899	6	47	4	2.53	563.33	ForwardHFNcog	Forward	Noncognate	High-FREQ	voiceless
211	slager	butcher	943	27.45	872	930	6	7	5	4.13	544.87	ForwardLFNcog	Forward	Noncognate	Low-FREQ	voiced
212	sleutel	key	745	26.92	860	897	7	34	6	4.87	547.03	ForwardHFNcog	Forward	Noncognate	High-FREQ	voiceless

Stimuli	Translation	Latency	cycles	**P.Latency.lm**	**P.Latency.zm**	**Stim.length**	freq	**Lev dist**	concreteness	Mean.N.RT	category	Direction	cognacy	freq.cat	Ph.Onset	voiceless
214	som	sum	1024	23.57	780	714	3	11	1	3.53	562.66	ForwardLFCog	Forward	Cognate	Low-FREQ	voiceless
215	spier	muscle	888	27.26	868	918	5	6	5	4.67	546.21	ForwardLFNcog	Forward	Noncognate	Low-FREQ	voiced
216	stad	city	823	25.74	831	828	4	249	4	3.73	562.74	ForwardHFNcog	Forward	Noncognate	High-FREQ	voiceless
217	stier	bull	1093	27.37	870	925	5	10	5	4.93	589.68	ForwardLFNcog	Forward	Noncognate	Low-FREQ	voiced
218	straat	street	728	22.47	754	662	6	147	2	2.78	605.54	ForwardHFCog	Forward	Cognate	High-FREQ	voiceless
219	tafel	table	825	25.88	835	836	5	189	3	4.93	509.55	ForwardHFCog	Forward	Cognate	High-FREQ	voiceless
220	tante	aunt	810	26.97	861	900	5	96	3	3.13	551.26	ForwardHFNcog	Forward	Noncognate	High-FREQ	vowel
221	thee	tea	741	27.12	864	909	4	51	2	4.86	524.28	ForwardHFCog	Forward	Cognate	High-FREQ	voiceless
222	trofee	trophy	835	27.49	873	933	6	1	3	4.13	589.21	ForwardLFCog	Forward	Cognate	Low-FREQ	voiceless
223	vader	father	755	25.34	822	805	5	546	3	3.13	544.72	ForwardHFCog	Forward	Cognate	High-FREQ	voiceless
224	veld	field	891	23.58	780	714	4	56	2	4.53	516.51	ForwardLFCog	Forward	Cognate	Low-FREQ	voiceless
225	verf	paint	880	27.27	868	919	4	26	5	4.8	537.92	ForwardLFNcog	Forward	Noncognate	Low-FREQ	voiceless
226	verhaal	story	769	26.09	840	848	7	161	7	2.87	547.03	ForwardHFNcog	Forward	Noncognate	High-FREQ	voiceless
227	verkeer	traffic	947	28.11	888	973	7	38	7	3.47	543.23	ForwardLFNcog	Forward	Noncognate	Low-FREQ	voiceless
228	vies	dirty	809	27.26	868	918	4	15	4	3	572.5	ForwardLF	Forw	Nonco	Low-	voice

											4	Ncog	ard	gnate	FREQ	d
Stimuli	Translation	Latency	cycles	**P.Latency.lm**	**P.Latency.zm**	**Stim.length**	freq	**Lev dist**	concreteness	Mean.N.RT	category	Direction	cognacy	freq.cat	Ph.Onset	voiceless
230	vloek	curse	837	27.42	872	928	5	8	5	2.13	550.56	ForwardLFNcog	Forward	Noncognate	Low-FREQ	voiceless
231	voet	foot	872	26.60	852	878	4	96	2	4.87	522.38	ForwardHFCog	Forward	Cognate	High-FREQ	voiceless
232	vos	fox	867	27.41	871	928	3	4	2	4.87	535.74	ForwardLFCog	Forward	Cognate	Low-FREQ	voiceless
233	vraag	question	754	25.17	818	796	5	456	8	2.47	564.08	ForwardHFNcog	Forward	Noncognate	High-FREQ	voiceless
234	vriend	friend	724	21.77	737	631	6	143	1	2.4	508.27	ForwardHFCog	Forward	Cognate	High-FREQ	voiceless
235	vrouw	woman	798	25.17	818	796	5	597	5	4.2	578.42	ForwardHFNcog	Forward	Noncognate	High-FREQ	voiced
236	warmte	warmth	753	22.51	754	664	6	47	1	3.6	533.33	ForwardLFCog	Forward	Cognate	Low-FREQ	voiced
237	wereld	world	707	21.60	733	624	6	443	2	3.33	555.67	ForwardHFCog	Forward	Cognate	High-FREQ	voiced
238	werk	work	683	21.13	722	605	4	495	1	3.2	546.85	ForwardHFCog	Forward	Cognate	High-FREQ	voiced
239	wezel	weasel	985	23.79	785	725	5	2	2	4.67	649.28	ForwardLFCog	Forward	Cognate	Low-FREQ	voiced
240	wijn	wine	680	26.59	852	877	4	140	2	4.8	548.9	ForwardHFCog	Forward	Cognate	High-FREQ	voiced
241	wolk	cloud	871	27.29	869	920	4	15	4	4.4	562.87	ForwardLFNcog	Forward	Noncognate	Low-FREQ	voiceless
242	woord	word	683	21.42	729	617	5	281	1	2.67	541.36	ForwardHFCog	Forward	Cognate	High-FREQ	voiced
243	zacht	soft	747	26.72	855	885	5	113	4	3.73	530.08	ForwardLFNcog	Forward	Noncognate	Low-FREQ	voiceless

Stimuli	Translation	Latency	cycles	**P.Latency.lm**	**P.Latency.zm**	**Stim.length**	freq	**Lev dist**	concreteness	Mean.N.RT	category	Direction	cognacy	freq.cat	Ph.Onset	voiceless
245	zilver	silver	745	22.58	756	667	6	12	1	4.67	567.55	ForwardLFCog	Forward	Cognate	Low-FREQ	voiceless
246	zoon	son	975	26.10	840	848	4	151	2	3.07	589.43	ForwardHFCog	Forward	Cognate	High-FREQ	voiceless
247	zorg	care	1092	26.74	855	886	4	82	3	2.8	528.31	ForwardHFNcog	Forward	Noncognate	High-FREQ	voiceless
248	zwart	black	693	26.20	842	854	5	81	4	4.27	552.74	ForwardHFNcog	Forward	Noncognate	High-FREQ	voiced
249	zweet	sweat	825	24.94	812	784	5	37	2	4.67	523.36	ForwardLFCog	Forward	Cognate	Low-FREQ	voiceless

11.4 **Descriptive Statistics**

The descriptive statistic tables below show the means, standard deviations, median, minimum, maximum, range, and standard error for latency, cycle-time, Z-scaled & L-scaled latency, frequency, and Levenshtein Distance, in each of the 11 conditions used for the correlational tests.

Global

N = 249	mean	SD	median	min	max	range	SE
Latency	828.36	118.45	810	621	1203	582	7.51
Cycle-time	25.61	2.04	26.25	21.13	30.77	9.63	0.13
Frequency	110.29	138.94	56	1	816	815	8.81
Levenshtein Distance	3.51	1.87	3	1	9	8	0.12
L-scaled Latency	828.36	48.62	843.79	721.67	951.42	229.74	3.08
Z-scaled Latency	828.05	111.77	857.28	604.65	1166.17	561.52	7.08

Forward Translation-direction

N = 124	mean	SD	median	min	max	range	SE
Latency	809.05	111.59	777.50	621.00	1093.00	472.00	10.02
Cycle-time	25.67	2.01	26.30	21.13	30.37	9.23	0.18
Frequency	107.02	138.25	50.00	1.00	597.00	596.00	12.42
Levenshtein Distance	3.45	1.81	3.00	1.00	8.00	7.00	0.16
L-scaled Latency	829.85	47.95	844.86	721.67	941.92	220.24	4.31
Z-scaled Latency	831.33	110.06	859.92	604.65	1134.92	530.27	9.88

Backward Translation-direction

N = 125	mean	SD	median	min	max	range	SE
Latency	847.51	122.33	828.00	631.00	1203.00	572.00	10.94
Cycle-time	25.54	2.07	26.25	21.18	30.77	9.59	0.19
Frequency	113.52	140.11	59.00	2.00	816.00	814.00	12.53
Levenshtein Distance	3.56	1.94	3.00	1.00	9.00	8.00	0.17
L-scaled Latency	826.87	49.42	843.79	722.66	951.42	228.76	4.42
Z-scaled Latency	824.80	113.80	857.28	606.35	1166.17	559.82	10.18

Forward high-frequency cognates

N = 32	mean	SD	median	min	max	range	SE
Latency	728.09	72.32	706.5	621	975	354	12.78
Cycle-time	24.11	2.03	23.58	21.13	27.12	5.99	0.36
Frequency	213.75	165.78	147.5	36	554	518	29.31
Levenshtein Distance	1.94	0.72	2	1	3	2	0.13
L-scaled Latency	792.65	48.41	780.02	721.67	864.45	142.77	8.56
Z-scaled Latency	747.58	103.49	714.42	604.65	909.45	304.8	18.3

Forward Low-frequency Cognate Condition

N = 31	mean	SD	median	min	max	range	SE
Latency	804.74	89.51	804	636	1024	388	16.08
Cycle-time	24.78	1.89	24.14	22.51	27.49	4.99	0.34
Frequency	21.29	20.77	12	1	91	90	3.73
Levenshtein Distance	1.97	0.71	2	1	3	2	0.13
L-scaled Latency	808.53	45.14	793.46	754.4	873.33	118.94	8.11
Z-scaled Latency	781.44	102.95	742.4	663.95	932.85	268.9	18.49

Forward High-frequency Noncognate Condition

N = 32	mean	SD	median	min	max	range	SE
Latency	802.53	93.13	773	673	1092	419	16.46
Cycle-time	26.54	1.12	26.44	25.17	30.37	5.2	0.2
Frequency	156.53	135.34	98.5	32	597	565	23.93
Levenshtein Distance	5.16	1.39	5	3	8	5	0.25
L-scaled Latency	850.55	26.72	848.25	817.93	941.92	123.99	4.72
Z-scaled Latency	876.62	72.25	868.3	796.19	1134.92	338.73	12.77

Forward Low-frequency Noncognate Condition

N = 29	mean	SD	median	min	max	Range	SE
Latency	910.17	112.13	892	663	1093	430	20.82
Cycle-time	27.39	0.62	27.36	25.65	29.62	3.97	0.11
Frequency	26.28	45.63	15	1	236	235	8.47
Levenshtein Distance	4.83	0.93	5	3	7	4	0.17
L-scaled Latency	870.87	14.71	870.12	829.49	924.08	94.59	2.73
Z-scaled Latency	927.09	40.01	924.31	822.95	1078.49	255.54	7.43

Backward High-frequency Cognate Condition

N = 32	mean	SD	median	min	max	Range	SE
Latency	738.69	71.55	727.5	631	974	343	12.65
Cycle-time	24.06	2.05	23.45	21.18	27.35	6.18	0.36
Frequency	223.09	192.18	189.5	21	816	795	33.97
Levenshtein Distance	2.09	0.82	2	1	4	3	0.14
L-scaled Latency	791.45	48.84	776.98	722.66	870.04	147.38	8.63
Z-scaled Latency	745.21	105.3	708.23	606.35	924.11	317.76	18.61

Backward High-frequency Cognate Condition

N = 32	mean	SD	median	min	max	Range	SE
Latency	738.69	71.55	727.5	631	974	343	12.65
Cycle-time	24.06	2.05	23.45	21.18	27.35	6.18	0.36
Frequency	223.09	192.18	189.5	21	816	795	33.97
Levenshtein Distance	2.09	0.82	2	1	4	3	0.14
L-scaled Latency	791.45	48.84	776.98	722.66	870.04	147.38	8.63
Z-scaled Latency	745.21	105.3	708.23	606.35	924.11	317.76	18.61

Backward High-frequency Cognate Condition

N = 32	mean	SD	median	min	max	Range	SE
Latency	738.69	71.55	727.5	631	974	343	12.65
Cycle-time	24.06	2.05	23.45	21.18	27.35	6.18	0.36
Frequency	223.09	192.18	189.5	21	816	795	33.97
Levenshtein Distance	2.09	0.82	2	1	4	3	0.14
L-scaled Latency	791.45	48.84	776.98	722.66	870.04	147.38	8.63
Z-scaled Latency	745.21	105.3	708.23	606.35	924.11	317.76	18.61

Backward Low-frequency Cognate Condition

N = 31	mean	SD	median	min	max	range	SE
Latency	830.32	84.42	844	711	1034	323	15.16
Cycle-time	24.39	1.85	23.78	22.12	27.49	5.37	0.33
Frequency	31.32	34.54	24	2	183	181	6.2
Levenshtein Distance	1.87	0.67	2	1	3	2	0.12
L-scaled Latency	799.27	44.24	784.86	745.15	873.34	128.19	7.95
Z-scaled Latency	760.88	100.55	724.36	646.62	932.87	286.25	18.06

Backward High-frequency Noncognate Condition

N = 32	mean	SD	median	min	max	range	SE
Latency	844.91	83.15	818.5	669	1039	370	14.7
Cycle-time	26.4	0.85	26.34	24.97	28.41	3.45	0.15
Frequency	164.47	101.68	171.5	15	403	388	17.97
Levenshtein Distance	5.13	1.41	5	2	8	6	0.25
L-scaled Latency	847.2	20.36	845.94	813.05	895.29	82.25	3.6
Z-scaled Latency	867.12	51.16	862.58	785.16	993.3	208.14	9.04

Backward Low Frequency Noncognate Condition

N = 30	mean	SD	median	min	max	range	SE
Latency	984.13	105.12	974.5	769	1203	434	19.19
Cycle-time	27.42	1.12	27.39	23.95	30.77	6.81	0.2
Frequency	27.23	33.68	14.5	2	164	162	6.15
Levenshtein Distance	5.2	1.35	5	3	9	6	0.25
L-scaled Latency	871.49	26.66	870.95	788.86	951.42	162.56	4.87
Z-scaled Latency	930.61	73.11	926.53	732.69	1166.17	433.48	13.35

11.5 **Spearman's Correlation Output**

P-values of "0" are, in reality, simply lower than the 2-decimal limit.

Latency * Frequency			
	estimate	statistic	p.value
Global	-0.49	3845228	0
Forward	-0.51	479212.3	0
Backward	-0.51	490508.3	0

ForwardHFCog	-0.05	5725.05	0.79
ForwardLFCog	-0.43	7079.06	0.02
ForwardHFNcog	-0.04	5699.07	0.81
ForwardLFNcog	-0.27	5160.63	0.15
BackwardHFCog	-0.38	7535.38	0.03
BackwardLFCog	-0.09	5431.19	0.61
BackwardHFNcog	-0.14	6201.71	0.46
BackwardLFNcog	-0.42	6385.9	0.02
Cycles * Latency			
	estimate	statistic	p.value
Global	0.47	1354395	0
Forward	0.43	182042.5	0
Backward	0.54	149705.1	0
ForwardHFCog	0.1	4897.9	0.58
ForwardLFCog	-0.11	5491.11	0.57
ForwardHFNcog	0.17	4521.91	0.35
ForwardLFNcog	0.42	2354	0.02
BackwardHFCog	0.09	4950.95	0.61
BackwardLFCog	0.16	4180	0.4
BackwardHFNcog	0.2	4339.8	0.26
BackwardLFNcog	0.21	3568.59	0.27
Cycles * frequency			
	estimate	statistic	p.value
Global	-0.48	3795609	0
Forward	-0.46	463053	0
Backward	-0.5	486933.9	0
ForwardHFCog	-0.38	7506	0.03
ForwardLFCog	-0.21	5989.49	0.26
ForwardHFNcog	-0.7	9298.7	0
ForwardLFNcog	-0.76	7127.54	0
BackwardHFCog	-0.6	8716.3	0
BackwardLFCog	-0.35	6704.7	0.05
BackwardHFNcog	-0.92	10481.46	0
BackwardLFNcog	-0.77	7964.82	0
Latency * LD			
	estimate	statistic	p.value
Global	0.43	1469309	0
Forward	0.33	212466.2	0
Backward	0.54	151167	0
ForwardHFCog	0.04	5213.09	0.81
ForwardLFCog	-0.35	6703.44	0.05

ForwardHFNcog	-0.01	5487.48	0.97
ForwardLFNcog	0.31	2785.27	0.1
BackwardHFCog	0.02	5335.51	0.9
BackwardLFCog	0.1	4474.73	0.6
BackwardHFNcog	0.46	2960.97	0.01
BackwardLFNcog	0.33	3025.79	0.08
Cycles * LD			
	estimate	statistic	p.value
Global	0.63	962546.6	0
Forward	0.61	124146.3	0
Backward	0.64	115903.1	0
ForwardHFCog	0.53	2545.99	0
ForwardLFCog	0.78	1088.53	0
ForwardHFNcog	-0.29	7014.3	0.11
ForwardLFNcog	-0.06	4314.52	0.75
BackwardHFCog	0.58	2268.72	0
BackwardLFCog	0.86	679.42	0
BackwardHFNcog	-0.12	6134.75	0.5
BackwardLFNcog	-0.25	5639.9	0.17
Frequency * LD			
	estimate	statistic	p.value
Global	0	2578018	0.98
Forward	0.03	306931.6	0.71
Backward	-0.04	338213.9	0.67
ForwardHFCog	0.27	3983.57	0.14
ForwardLFCog	0.1	4461.31	0.59
ForwardHFNcog	0.23	4203.97	0.21
ForwardLFNcog	0.38	2505.74	0.04
BackwardHFCog	-0.18	6448.97	0.32
BackwardLFCog	-0.17	5792.86	0.37
BackwardHFNcog	0.17	4510.58	0.34
BackwardLFNcog	0.06	4203.59	0.73
Cycles * L-scaled			
	estimate	statistic	p.value
Global	1	0	0
Forward	1	0	0
Backward	1	0	0
ForwardHFCog	1	0	0
ForwardLFCog	1	0	0
ForwardHFNcog	1	0	0
ForwardLFNcog	1	9.02E-13	0

BackwardHFCog	1	0	0
BackwardLFCog	1	0	0
BackwardHFNcog	1	0	0
BackwardLFNcog	1	0	0
Frequency * L-scaled			
	estimate	statistic	p.value
Global	-0.48	3795609	0
Forward	-0.46	463053	0
Backward	-0.5	486933.9	0
ForwardHFCog	-0.38	7506	0.034813
ForwardLFCog	-0.21	5989.49	0.262541
ForwardHFNcog	-0.7	9298.7	0.000007
ForwardLFNcog	-0.76	7127.54	0.000002
BackwardHFCog	-0.6	8716.3	0.000305
BackwardLFCog	-0.35	6704.7	0.052315
BackwardHFNcog	-0.92	10481.46	0
BackwardLFNcog	-0.77	7964.82	0.000001
Latency * L-scaled			
	estimate	statistic	p.value
Global	0.47	1354395	0
Forward	0.43	182042.5	0.000001
Backward	0.54	149705.1	0
ForwardHFCog	0.1	4897.9	0.577465
ForwardLFCog	-0.11	5491.11	0.566416
ForwardHFNcog	0.17	4521.91	0.348819
ForwardLFNcog	0.42	2354	0.024108
BackwardHFCog	0.09	4950.95	0.614339
BackwardLFCog	0.16	4180	0.396596
BackwardHFNcog	0.2	4339.8	0.26136
BackwardLFNcog	0.21	3568.59	0.274541
LD * L-scaled			
	estimate	statistic	p.value
Global	0.63	962546.6	0
Forward	0.61	124146.3	0
Backward	0.64	115903.1	0
ForwardHFCog	0.53	2545.99	0.00167
ForwardLFCog	0.78	1088.53	0
ForwardHFNcog	-0.29	7014.3	0.113057
ForwardLFNcog	-0.06	4314.52	0.746644
BackwardHFCog	0.58	2268.72	0.000447
BackwardLFCog	0.86	679.42	0

BackwardHFNcog	-0.12	6134.75	0.497528
BackwardLFNcog	-0.25	5639.9	0.174359
Cycles * Z-scaled			
	estimate	statistic	p.value
Global	1	0	0
Forward	1	0	0
Backward	1	0	0
ForwardHFCog	1	0	0
ForwardLFCog	1	0	0
ForwardHFNcog	1	0	0
ForwardLFNcog	1	0	0
BackwardHFCog	1	0	0
BackwardLFCog	1	0	0
BackwardHFNcog	1	0	0
BackwardLFNcog	1	0	0
Frequency * Z-scaled			
	estimate	statistic	p.value
Global	-0.48	3795609	0
Forward	-0.46	463053	0
Backward	-0.5	486933.9	0
ForwardHFCog	-0.38	7506	0.034813
ForwardLFCog	-0.21	5989.49	0.262541
ForwardHFNcog	-0.7	9298.7	0.000007
ForwardLFNcog	-0.76	7127.54	0.000002
BackwardHFCog	-0.6	8716.3	0.000305
BackwardLFCog	-0.35	6704.7	0.052315
BackwardHFNcog	-0.92	10481.46	0
BackwardLFNcog	-0.77	7964.82	0.000001
Latency * Z-scaled			
	estimate	statistic	p.value
Global	0.47	1354395	0
Forward	0.43	182042.5	0.000001
Backward	0.54	149705.1	0
ForwardHFCog	0.1	4897.9	0.577465
ForwardLFCog	-0.11	5491.11	0.566416
ForwardHFNcog	0.17	4521.91	0.348819
ForwardLFNcog	0.42	2354	0.024108
BackwardHFCog	0.09	4950.95	0.614339
BackwardLFCog	0.16	4180	0.396596
BackwardHFNcog	0.2	4339.8	0.26136
BackwardLFNcog	0.21	3568.59	0.274541

Z-scaled * LD			
	estimate	statistic	p.value
Global	0.63	962546.6	0
Forward	0.61	124146.3	0
Backward	0.64	115903.1	0
ForwardHFCog	0.53	2545.99	0.00167
ForwardLFCog	0.78	1088.53	0
ForwardHFNcog	-0.29	7014.3	0.113057
ForwardLFNcog	-0.06	4314.52	0.746644
BackwardHFCog	0.58	2268.72	0.000447
BackwardLFCog	0.86	679.42	0
BackwardHFNcog	-0.12	6134.75	0.497528
BackwardLFNcog	-0.25	5639.9	0.174359
Mean-regressed			
	estimate	statistic	p.value
Latency * Cycle-time	0.809524	16	0.021776
Latency * Frequency	-0.61905	136	0.11498
Frequency * Cycle-time	-0.61905	136	0.11498
LD * Cycle-time	0.761905	20	0.036756
Frequency * LD	-0.09524	92	0.840129
Latency * LD	0.5	42	0.216171

11.6 **Analysis Of Variance (ANOVA) Output**

Latency						
Interaction	DF	Sum Of Squares	Mean Square	F value	P-value (>F)	Partial Omega-squared
Direction	1	92094.25	92094.25	11.49586	0.000815	0.040447
Cognacy	1	729728	729728	91.08983	1.62E-18	0.265681
Frequency Category	1	668639	668639	83.46427	2.74E-17	0.248788
Direction*Cognacy	1	24553.24	24553.24	3.064909	0.081272	0.008225
Direction*Frequency Category	1	8479.026	8479.026	1.058412	0.304609	0.000235
Cognacy*Frequency Category	1	24044.33	24044.33	3.001384	0.084473	0.007974
Direction*Cognacy*Frequency Category	1	1070.708	1070.708	0.133653	0.714994	0.003491
Residuals	241	1930671	8011.081			
Cycle-time						
Interaction	DF	Sum Of Squares	Mean Square	F value	P-value (>F)	Partial Omega-squared

Direction	1	0.97269	0.97269	0.406576	0.524318	0.002389
Cognacy	1	416.5057	416.5057	174.0959	2.77E-30	0.410087
Frequency Category	1	31.69298	31.69298	13.2474	0.000334	0.04688
Direction*Cognacy	1	0.3873	0.3873	0.161888	0.687781	0.003377
Direction*Frequency Category	1	0.116917	0.116917	0.048871	0.825228	0.003834
Cognacy*Frequency Category	1	3.000527	3.000527	1.254195	0.263868	0.00102
Direction*Cognacy*Frequency Category	1	0.990109	0.990109	0.413857	0.52063	0.00236
Residuals	241	576.5666	2.392392			
L-scaled Latency						
Interaction	DF	Sum Of Squares	Mean Square	F value	P-value (>F)	Partial Omega-squared
Direction	1	553.4622	553.4622	0.406576	0.524318	0.002389
Cognacy	1	236992.5	236992.5	174.0959	2.77E-30	0.410087
Frequency Category	1	18033.37	18033.37	13.2474	0.000334	0.04688
Direction*Cognacy	1	220.3744	220.3744	0.161888	0.687781	0.003377
Direction*Frequency Category	1	66.52624	66.52624	0.048871	0.825228	0.003834
Cognacy*Frequency Category	1	1707.305	1707.305	1.254195	0.263868	0.00102
Direction*Cognacy*Frequency Category	1	563.3737	563.3737	0.413857	0.52063	0.00236
Residuals	241	328067.4	1361.276			
Z-scaled Latency						
Interaction	DF	Sum Of Squares	Mean Square	F value	P-value (>F)	Partial Omega-squared
Direction	1	2655.965	2655.965	0.367894	0.544726	0.002545
Cognacy	1	1231477	1231477	170.5793	7.77E-30	0.405131
Frequency Category	1	103091.2	103091.2	14.27978	0.000199	0.050632
Direction*Cognacy	1	1046.638	1046.638	0.144976	0.703719	0.003446
Direction*Frequency Category	1	108.2496	108.2496	0.014994	0.902644	0.003972
Cognacy*Frequency Category	1	16197.92	16197.92	2.243671	0.135471	0.00497
Direction*Cognacy*Frequency Category	1	3783.462	3783.462	0.52407	0.469813	0.001915
Residuals	241	1739871	7219.381			

11.7 **Generalized Additive Regression Model Output**

Latency																
GAM output						**X2 analysis of deviance**					**GAM ANOVA test**					
	Estimate	SE	Z Value	Pr(>|z|)			DF	X2	p-value		term	DF	Sum Of Squares	Mean Square	F Value	p value
(Intercept)	700.73	170.28	4.12	0.00		Direction	1.00	0.13	0.72		Direction	1.00	92094.25	92094.25	12.20	0.00
Direction: Forward	-36.02	100.63	-0.36	0.72		Cognacy	1.00	0.16	0.69		Cognacy	1.00	729728.00	729728.00	96.67	0.00
Cognacy: Noncognate	40.75	101.60	0.40	0.69		Frequency-category	1.00	2.95	0.09		Frequency-category	1.00	668639.04	668639.04	88.58	0.00
Frequency-category: Low-frequency	176.39	102.73	1.72	0.09		Phonetic Onset	2.00	6.38	0.04		Phonetic Onset	2.00	21288.98	10644.49	1.41	0.25
Phonetic Onset: voiceless	244.42	104.17	2.35	0.02		Stimulus Length	1.00	0.02	0.90		Stimulus Length	1.00	181.45	181.45	0.02	0.88
Phonetic Onset: vowel	-65.10	234.96	-0.28	0.78		Concreteness	1.00	0.04	0.84		Concreteness	1.00	7.83	7.83	0.00	0.97
Stimulus Length	4.02	30.69	0.13	0.90		Direction*Cognacy	1.00	4.53	0.03		Direction *Cognacy	1.00	24175.03	24175.03	3.20	0.07
Concreteness	-7.34	37.23	-0.20	0.84		Direction*Frequency-	1.00	1.17	0.28		Direction *Frequency-category	1.00	9505.61	9505.61	1.26	0.26

						categor y									
Direction: Forward*Cognac y: Noncognate	-52.80	24.80	-2.13	0.03		Directi on*Ph onetic Onset	2.00	2.46	0.29		Direction *Phoneti c Onset	2.00	83363.22	41681.61	5.52	0.00
Direction: Forward*Freque ncy-category: Low-frequency	-27.02	24.99	-1.08	0.28		Directi on*Sti mulus Length	1.00	0.42	0.52		Direction *Stimulu s Length	1.00	14372.87	14372.87	1.90	0.17
Direction: Forward*Phoneti c Onset: voiceless	34.87	25.54	1.37	0.17		Directi on*Co ncrete ness	1.00	0.43	0.51		Direction *Concret eness	1.00	1079.75	1079.75	0.14	0.71
Direction: Forward*Phoneti c Onset: vowel	67.34	68.07	0.99	0.32		Cognac y*Freq uency- categor y	1.00	1.83	0.18		Cognacy* Frequenc y- category	1.00	13823.05	13823.05	1.83	0.18
Direction: Forward*Stimulu s Length	-7.52	11.60	-0.65	0.52		Cognac y*Phon etic Onset	2.00	1.74	0.42		Cognacy* Phonetic Onset	2.00	18765.23	9382.62	1.24	0.29
Direction: Forward*Concret eness	10.50	16.01	0.66	0.51		Cognac y*Stim ulus Length	1.00	0.23	0.63		Cognacy* Stimulus Length	1.00	5601.55	5601.55	0.74	0.39
Cognacy: Noncognate*Fre quency-category: Low-frequency	33.59	24.81	1.35	0.18		Cognac y*Conc retenes s	1.00	0.66	0.42		Cognacy* Concrete ness	1.00	428.68	428.68	0.06	0.81
Cognacy: Noncognate*Pho	-14.95	25.13	-0.60	0.55		Freque ncy-	2.00	3.84	0.15		Frequenc y-	2.00	26981.36	13490.68	1.79	0.17

netic Onset: voiceless						category*Phonetic Onset					category*Phonetic Onset					
Cognacy: Noncognate*Phonetic Onset: vowel	76.72	72.53	1.06	0.29		Frequency-category*Stimulus Length	1.00	1.74	0.19		Frequency-category*Stimulus Length	1.00	23335.27	23335.27	3.09	0.08
Cognacy: Noncognate*Stimulus Length	5.58	11.74	0.48	0.63		Frequency-category*Concreteness	1.00	0.05	0.83		Frequency-category*Concreteness	1.00	4.23	4.23	0.00	0.98
Cognacy: Noncognate*Concreteness	12.67	15.56	0.81	0.42		Phonetic Onset*Stimulus Length	2.00	4.86	0.09		Phonetic Onset*Stimulus Length	2.00	46060.40	23030.20	3.05	0.05
Frequency-category: Low-frequency*Phonetic Onset: voiceless	-34.46	25.00	-1.38	0.17		Phonetic Onset*Concreteness	2.00	3.72	0.16		Phonetic Onset*Concreteness	2.00	30853.79	15426.90	2.04	0.13
Frequency-category: Low-frequency*Phonetic Onset: vowel	77.44	69.94	1.11	0.27		Stimulus Length*Concreteness	1.00	0.09	0.76		Stimulus Length*Concreteness	1.00	761.22	761.22	0.10	0.75

Frequency-category: Low-frequency*Stimulus Length	-15.86	12.01	-1.32	0.19							Residuals	221.00	1668228.37	7548.54		
Frequency-category: Low-frequency*Concreteness	3.34	15.43	0.22	0.83												
Phonetic Onset: voiceless*Stimulus Length	-18.58	11.86	-1.57	0.12												
Phonetic Onset: vowel*Stimulus Length	28.54	25.21	1.13	0.26												
Phonetic Onset: voiceless*Concreteness	-30.06	16.20	-1.85	0.06												
Phonetic Onset: vowel*Concreteness	-41.72	41.31	-1.01	0.31												
Stimulus Length*Concreteness	2.03	6.78	0.30	0.76												
N = 249	Model parameters = 28	Deviance explained = 0.521	R2-adj = 0.46	UBRE\AIC = 111.36												

Cycle-time																
GAM output						**X2 analysis of deviance**					**GAM ANOVA test**					
	Estimate	SE	Z Value	Pr(>|z|)			DF	X2	p-value		term	DF	Sum Of Squares	Mean Square	F Value	p value
(Intercept)	26.93	2.61	10.31	0.00		Direction	1.00	0.66	0.42		Direction	1.00	0.97	0.97	0.45	0.50
Direction: Forward	-1.25	1.54	-0.81	0.42		Cognacy	1.00	0.03	0.85		Cognacy	1.00	416.51	416.51	192.64	0.00
Cognacy: Noncognate	0.29	1.56	0.19	0.85		Frequency-category	1.00	0.49	0.48		Frequency-category	1.00	31.69	31.69	14.66	0.00
Frequency-category: Low-frequency	-1.11	1.58	-0.70	0.48		Phonetic Onset	2.00	0.54	0.76		Phonetic Onset	2.00	2.05	1.02	0.47	0.62
Phonetic Onset: voiceless	-1.07	1.60	-0.67	0.50		Stimulus Length	1.00	2.20	0.14		Stimulus Length	1.00	29.85	29.85	13.81	0.00
Phonetic Onset: vowel	0.44	3.60	0.12	0.90		Concreteness	1.00	0.12	0.73		Concreteness	1.00	9.08	9.08	4.20	0.04
Stimulus Length	-0.70	0.47	-1.48	0.14		Direction*Cognacy	1.00	0.00	0.96		Direction *Cognacy	1.00	0.04	0.04	0.02	0.89
Concreteness	0.20	0.57	0.34	0.73		Direction*Frequency-category	1.00	0.01	0.93		Direction *Frequency-category	1.00	0.19	0.19	0.09	0.77
Direction:	-0.02	0.38	-0.05	0.96		Directi	2.00	2.23	0.33		Direction	2.00	7.44	3.72	1.72	0.18

Forward*Cognacy: Noncognate						on*Phonetic Onset					*Phonetic Onset					
Direction: Forward*Frequency-category: Low-frequency	0.03	0.38	0.09	0.93		Direction*Stimulus Length	1.00	0.40	0.53		Direction *Stimulus Length	1.00	0.51	0.51	0.23	0.63
Direction: Forward*Phonetic Onset: voiceless	0.58	0.39	1.49	0.14		Direction*Concreteness	1.00	0.27	0.61		Direction *Concreteness	1.00	0.16	0.16	0.08	0.78
Direction: Forward*Phonetic Onset: vowel	0.12	1.04	0.11	0.91		Cognacy*Frequency-category	1.00	1.19	0.27		Cognacy* Frequency-category	1.00	2.45	2.45	1.13	0.29
Direction: Forward*Stimulus Length	0.11	0.18	0.63	0.53		Cognacy*Phonetic Onset	2.00	1.16	0.56		Cognacy* Phonetic Onset	2.00	3.42	1.71	0.79	0.45
Direction: Forward*Concreteness	0.13	0.25	0.51	0.61		Cognacy*Stimulus Length	1.00	12.27	0.00		Cognacy* Stimulus Length	1.00	33.42	33.42	15.46	0.00
Cognacy: Noncognate*Frequency-category: Low-frequency	0.42	0.38	1.09	0.27		Cognacy*Concreteness	1.00	0.64	0.42		Cognacy* Concreteness	1.00	1.80	1.80	0.83	0.36
Cognacy: Noncognate*Phonetic Onset: voiceless	-0.41	0.39	-1.07	0.28		Frequency-category*Phonetic	2.00	3.25	0.20		Frequency-category *Phonetic Onset	2.00	5.61	2.81	1.30	0.28

						Onset									
Cognacy: Noncognate*Phonetic Onset: vowel	-0.32	1.11	-0.29	0.77		Frequency-category*Stimulus Length	1.00	0.37	0.54		Frequency-category*Stimulus Length	1.00	0.27	0.27	0.12	0.73
Cognacy: Noncognate*Stimulus Length	0.63	0.18	3.50	0.00		Frequency-category*Concreteness	1.00	1.83	0.18		Frequency-category*Concreteness	1.00	3.99	3.99	1.85	0.18
Cognacy: Noncognate*Concreteness	-0.19	0.24	-0.80	0.42		Phonetic Onset*Stimulus Length	2.00	1.42	0.49		Phonetic Onset*Stimulus Length	2.00	2.62	1.31	0.61	0.55
Frequency-category: Low-frequency*Phonetic Onset: voiceless	-0.64	0.38	-1.67	0.10		Phonetic Onset*Concreteness	2.00	0.15	0.93		Phonetic Onset*Concreteness	2.00	0.26	0.13	0.06	0.94
Frequency-category: Low-frequency*Phonetic Onset: vowel	0.38	1.07	0.36	0.72		Stimulus Length*Concreteness	1.00	0.04	0.84		Stimulus Length*Concreteness	1.00	0.08	0.08	0.04	0.85
Frequency-category: Low-frequency*Stimul	0.11	0.18	0.61	0.54							Residuals	221.00	477.83	2.16		

us Length																
Frequency-category: Low-frequency*Concreteness	0.32	0.24	1.35	0.18												
Phonetic Onset: voiceless*Stimulus Length	0.21	0.18	1.16	0.25												
Phonetic Onset: vowel*Stimulus Length	0.00	0.39	-0.01	0.99												
Phonetic Onset: voiceless*Concreteness	0.08	0.25	0.32	0.75												
Phonetic Onset: vowel*Concreteness	-0.08	0.63	-0.12	0.91												
Stimulus Length*Concreteness	-0.02	0.10	-0.20	0.84												
N = 249	Model parameters = 28	Deviance explained = 0.536	R2-adj = 0.48	UBRE\AIC = 0.36879												
L-scaled latency																
GAM output						**X2 analysis of deviance**					**GAM ANOVA test**					

	Estimate	SE	Z Value	Pr(>|z|)			DF	X2	p-value		term	DF	Sum Of Squares	Mean Square	F Value	p value
(Intercept)	859.93	69.11	12.44	0.00		Direction	1.00	0.53	0.46		Direction	1.00	553.46	553.46	0.45	0.50
Direction: Forward	-29.84	40.84	-0.73	0.46		Cognacy	1.00	0.03	0.87		Cognacy	1.00	236992.50	236992.50	192.64	0.00
Cognacy: Noncognate	6.90	41.23	0.17	0.87		Frequency-category	1.00	0.40	0.53		Frequency-category	1.00	18033.37	18033.37	14.66	0.00
Frequency-category: Low-frequency	-26.36	41.69	-0.63	0.53		Phonetic Onset	2.00	0.44	0.80		Phonetic Onset	2.00	1163.72	581.86	0.47	0.62
Phonetic Onset: voiceless	-25.41	42.28	-0.60	0.55		Stimulus Length	1.00	1.79	0.18		Stimulus Length	1.00	16987.48	16987.48	13.81	0.00
Phonetic Onset: vowel	10.40	95.36	0.11	0.91		Concreteness	1.00	0.10	0.76		Concreteness	1.00	5166.67	5166.67	4.20	0.04
Stimulus Length	-16.64	12.45	-1.34	0.18		Direction*Cognacy	1.00	0.00	0.96		Direction*Cognacy	1.00	24.44	24.44	0.02	0.89
Concreteness	4.67	15.11	0.31	0.76		Direction*Frequency-category	1.00	0.01	0.93		Direction*Frequency-category	1.00	105.42	105.42	0.09	0.77
Direction: Forward*Cognacy: Noncognate	-0.48	10.06	-0.05	0.96		Direction*Phonetic Onset	2.00	1.82	0.40		Direction*Phonetic Onset	2.00	4230.86	2115.43	1.72	0.18
Direction:	0.83	10.14	0.08	0.93		Directi	1.00	0.32	0.57		Direction	1.00	288.13	288.13	0.23	0.63

Forward*Frequency-category: Low-frequency						on*Stimulus Length					*Stimulus Length					
Direction: Forward*Phonetic Onset: voiceless	13.91	10.37	1.34	0.18		Direction*Concreteness	1.00	0.22	0.64		Direction*Concreteness	1.00	93.50	93.50	0.08	0.78
Direction: Forward*Phonetic Onset: vowel	2.76	27.62	0.10	0.92		Cognacy*Frequency-category	1.00	0.97	0.32		Cognacy*Frequency-category	1.00	1391.52	1391.52	1.13	0.29
Direction: Forward*Stimulus Length	2.68	4.71	0.57	0.57		Cognacy*Phonetic Onset	2.00	0.94	0.62		Cognacy*Phonetic Onset	2.00	1945.70	972.85	0.79	0.45
Direction: Forward*Concreteness	3.02	6.50	0.46	0.64		Cognacy*Stimulus Length	1.00	9.98	0.00		Cognacy*Stimulus Length	1.00	19014.92	19014.92	15.46	0.00
Cognacy: Noncognate*Frequency-category: Low-frequency	9.91	10.07	0.98	0.32		Cognacy*Concreteness	1.00	0.52	0.47		Cognacy*Concreteness	1.00	1024.75	1024.75	0.83	0.36
Cognacy: Noncognate*Phonetic Onset: voiceless	-9.85	10.20	-0.97	0.33		Frequency-category*Phonetic Onset	2.00	2.64	0.27		Frequency-category*Phonetic Onset	2.00	3193.89	1596.95	1.30	0.28
Cognacy: Noncognate*Phonetic Onset:	-7.70	29.44	-0.26	0.79		Frequency-categor	1.00	0.30	0.58		Frequency-category	1.00	151.10	151.10	0.12	0.73

vowel						y*Stimulus Length					*Stimulus Length					
Cognacy: Noncognate*Stimulus Length	15.05	4.76	3.16	0.00		Frequency-category*Concreteness	1.00	1.49	0.22		Frequency-category*Concreteness	1.00	2273.03	2273.03	1.85	0.18
Cognacy: Noncognate*Concreteness	-4.55	6.32	-0.72	0.47		Phonetic Onset*Stimulus Length	2.00	1.15	0.56		Phonetic Onset*Stimulus Length	2.00	1490.36	745.18	0.61	0.55
Frequency-category: Low-frequency*Phonetic Onset: voiceless	-15.26	10.15	-1.50	0.13		Phonetic Onset*Concreteness	2.00	0.12	0.94		Phonetic Onset*Concreteness	2.00	148.59	74.30	0.06	0.94
Frequency-category: Low-frequency*Phonetic Onset: vowel	9.14	28.38	0.32	0.75		Stimulus Length*Concreteness	1.00	0.03	0.86		Stimulus Length*Concreteness	1.00	45.95	45.95	0.04	0.85
Frequency-category: Low-frequency*Stimulus Length	2.68	4.87	0.55	0.58							Residuals	221.00	271884.94	1230.25		
Frequency-category: Low-frequency*Concr	7.64	6.26	1.22	0.22												

eteness																
Phonetic Onset: voiceless*Stimulus Length	5.02	4.82	1.04	0.30												
Phonetic Onset: vowel*Stimulus Length	-0.06	10.23	-0.01	1.00												
Phonetic Onset: voiceless*Concreteness	1.90	6.58	0.29	0.77												
Phonetic Onset: vowel*Concreteness	-1.79	16.77	-0.11	0.91												
Stimulus Length*Concreteness	-0.50	2.75	-0.18	0.86												
N = 249	Model parameters = 28	Deviance explained = 0.536	R2-adj = 0.48	UBRE\AIC = 6.7668												
Z-scaled latency																
GAM output						**X2 analysis of deviance**					**GAM ANOVA test**					
	Estimate	SE	Z Value	Pr(>|z|)			DF	X2	p-value		term	DF	Sum Of Squares	Mean Square	F Value	p value
(Intercept)	891.70	159.9	5.57	0.00		Directi	1.00	0.59	0.44		Direction	1.00	2655.96	2655.96	0.40	0.53

		5				on										
Direction: Forward	-72.78	94.53	-0.77	0.44		Cognacy	1.00	0.01	0.93		Cognacy	1.00	1231477.09	1231477.09	187.25	0.00
Cognacy: Noncognate	8.81	95.44	0.09	0.93		Frequency-category	1.00	0.30	0.58		Frequency-category	1.00	103091.19	103091.19	15.68	0.00
Frequency-category: Low-frequency	-53.05	96.49	-0.55	0.58		Phonetic Onset	2.00	0.57	0.75		Phonetic Onset	2.00	3675.96	1837.98	0.28	0.76
Phonetic Onset: voiceless	-66.49	97.85	-0.68	0.50		Stimulus Length	1.00	1.48	0.22		Stimulus Length	1.00	84166.66	84166.66	12.80	0.00
Phonetic Onset: vowel	30.86	220.71	0.14	0.89		Concreteness	1.00	0.12	0.73		Concreteness	1.00	28058.13	28058.13	4.27	0.04
Stimulus Length	-35.05	28.83	-1.22	0.22		Direction*Cognacy	1.00	0.00	0.95		Direction*Cognacy	1.00	103.12	103.12	0.02	0.90
Concreteness	12.06	34.97	0.34	0.73		Direction*Frequency-category	1.00	0.00	0.99		Direction*Frequency-category	1.00	948.82	948.82	0.14	0.70
Direction: Forward*Cognacy: Noncognate	-1.51	23.30	-0.06	0.95		Direction*Phonetic Onset	2.00	1.86	0.39		Direction*Phonetic Onset	2.00	23851.80	11925.90	1.81	0.17
Direction: Forward*Frequency-category: Low-frequency	-0.18	23.48	-0.01	0.99		Direction*Stimulus Length	1.00	0.44	0.51		Direction*Stimulus Length	1.00	2207.33	2207.33	0.34	0.56

Direction: Forward*Phonetic Onset: voiceless	32.64	23.99	1.36	0.17		Direction*Concreteness	1.00	0.21	0.65		Direction*Concreteness	1.00	305.70	305.70	0.05	0.83
Direction: Forward*Phonetic Onset: vowel	8.06	63.94	0.13	0.90		Cognacy*Frequency-category	1.00	1.83	0.18		Cognacy*Frequency-category	1.00	14072.57	14072.57	2.14	0.14
Direction: Forward*Stimulus Length	7.21	10.90	0.66	0.51		Cognacy*Phonetic Onset	2.00	0.88	0.64		Cognacy*Phonetic Onset	2.00	10546.24	5273.12	0.80	0.45
Direction: Forward*Concreteness	6.86	15.04	0.46	0.65		Cognacy*Stimulus Length	1.00	9.20	0.00		Cognacy*Stimulus Length	1.00	94298.29	94298.29	14.34	0.00
Cognacy: Noncognate*Frequency-category: Low-frequency	31.51	23.30	1.35	0.18		Cognacy*Concreteness	1.00	0.35	0.55		Cognacy*Concreteness	1.00	4182.01	4182.01	0.64	0.43
Cognacy: Noncognate*Phonetic Onset: voiceless	-22.04	23.60	-0.93	0.35		Frequency-category*Phonetic Onset	2.00	2.62	0.27		Frequency-category*Phonetic Onset	2.00	17050.59	8525.29	1.30	0.28
Cognacy: Noncognate*Phonetic Onset: vowel	-15.79	68.13	-0.23	0.82		Frequency-category*Stimulus Length	1.00	0.10	0.75		Frequency-category*Stimulus Length	1.00	59.99	59.99	0.01	0.92

Cognacy: Noncognate*Stimulus Length	33.45	11.03	3.03	0.00		Frequency-category*Concreteness	1.00	1.61	0.20		Frequency-category*Concreteness	1.00	13021.48	13021.48	1.98	0.16
Cognacy: Noncognate*Concreteness	-8.70	14.62	-0.60	0.55		Phonetic Onset*Stimulus Length	2.00	1.37	0.50		Phonetic Onset*Stimulus Length	2.00	9618.31	4809.16	0.73	0.48
Frequency-category: Low-frequency*Phonetic Onset: voiceless	-34.90	23.48	-1.49	0.14		Phonetic Onset*Concreteness	2.00	0.14	0.93		Phonetic Onset*Concreteness	2.00	903.88	451.94	0.07	0.93
Frequency-category: Low-frequency*Phonetic Onset: vowel	22.71	65.69	0.35	0.73		Stimulus Length*Concreteness	1.00	0.06	0.80		Stimulus Length*Concreteness	1.00	480.19	480.19	0.07	0.79
Frequency-category: Low-frequency*Stimulus Length	3.61	11.28	0.32	0.75							Residuals	221.00	1453456.10	6576.72		
Frequency-category: Low-frequency*Concreteness	18.40	14.49	1.27	0.20												
Phonetic Onset: voiceless*Stimulu	12.55	11.14	1.13	0.26												

s Length																
Phonetic Onset: vowel*Stimulus Length	-1.11	23.68	-0.05	0.96												
Phonetic Onset: voiceless*Concreteness	4.32	15.22	0.28	0.78												
Phonetic Onset: vowel*Concreteness	-6.14	38.80	-0.16	0.87												
Stimulus Length*Concreteness	-1.61	6.37	-0.25	0.80												
N = 249	Model parameters = 28	Deviance explained = 0.531	R2-adj = 0.47	UBRE\AIC = 23.92												

11.8 **Statistical Script**

This statistical script was used to generate the results of this thesis in the program R (R Core Team, 2015). Acknowledgements for its creation and length must be given to Dr. Sean Roberts most of all, and also to Dr. Francisco Torreira, Merel Maslowski, and Jeremy Collins. All 4 helped troubleshoot issues.

```
library(lme4)
library(seewave)
library(arm)
library(HH)
library(Hmisc)
library(plyr)
library(psych)
library(gplots)
library(ggplot2)
library(lsr)
library(distr)
library(dbEmpLikeGOF)
library(entropy)
library(reshape2)
library(distrEx)
library(corrgram)
library(GoFKernel)
library(broom)
library(stringdist)
library(data.table)
library(gridExtra)
library(mgcv)
library(gamlss)
library(gamm4)


###########################################################################################
###########
###### Script created with R 3.2.2 14-08-2015 "Fire Safety" and Rstudio 0.99.473 ###########
##### if you don't have these packages, use the install.packages(PackageName) command ######
```

```
#####################################################################################
###########
# files will appear wherever current "source" is located
# import 'Multilink.csv' first, then run all.

Multilink <- read.table(file='Multilink.csv', header=TRUE, sep=',', dec='.')

############ Partial Omega Squared Function ################
# just highlight and source this.

# Partial omega-squared
partialOmegas <- function(mod){
  aovMod <- mod
  if(!any(class(aovMod) %in% 'aov')) aovMod <- aov(mod)
  sumAov     <- summary(aovMod)[[1]]
  residRow   <- nrow(sumAov)
  dfError    <- sumAov[residRow,1]
  msError    <- sumAov[residRow,3]
  nTotal     <- nrow(model.frame(aovMod))
  dfEffects  <- sumAov[1:{residRow-1},1]
  ssEffects  <- sumAov[1:{residRow-1},2]
  msEffects  <- sumAov[1:{residRow-1},3]
  partOmegas <- abs((dfEffects*(msEffects-msError)) /
                    (ssEffects + (nTotal -dfEffects)*msError))
  names(partOmegas) <- rownames(sumAov)[1:{residRow-1}]
  partOmegas
}

################################################################
################################################################

############# Levenshtein Distance calculation ################

# uses package "stringdistr"
Multilink$Levdist = stringdist(Multilink$Stimuli, Multilink$Translation, method= "lv")

# uses package "plyr"
# cognate vs noncog ; all numbers should match
```

```
count(Multilink$cognacy)
count(Multilink$Levdist)
count(Multilink$Levdist <= 3)


############################################################
############################################################

############### Linear Model, And Cycle->Latency Scaling ##################
###########################################################################
# generates 2 new data column that extrapolate possible latencies from models of cycle-time
# 2 models present: a simple linear model (good fit), and a Z-score model (better-than-good fit)
# used for the principal tests of the results.

## Linear Models ##
# just a check
#"intercept" seems to be the backward high-frequency cognate condition
summary(lm(Latency ~ cycles, data= Multilink))
summary(lm(cycles ~ Direction*cognacy*freq.cat, data= Multilink))
summary(lm(Latency ~ Direction*cognacy*freq.cat, data= Multilink))
summary(lm(Latency ~ freq.cat, data= Multilink))


## Linear model ##
# highlight the whole thing to run.

linmodlatency = lm(Latency ~ cycles, data= Multilink)
summary(linmodlatency)
Multilink$P.Latency.lm = predict(linmodlatency)


## Z-score model ##
#this has a better fit than the linear model.

Multilink$Latency.z = (Multilink$Latency - mean(Multilink$Latency)) / sd(Multilink$Latency)
Multilink$Latency.log = log10(Multilink$Latency)
Multilink$Latency.log.z = (Multilink$Latency.log - mean(Multilink$Latency.log)) /
sd(Multilink$Latency.log)
```

```
Multilink$cycles.z = (Multilink$cycles - mean(Multilink$cycles))  /sd(Multilink$cycles)
Multilink$Latency.approx.log = (Multilink$cycles.z * sd(Multilink$Latency.log)) +
mean(Multilink$Latency.log)
Multilink$P.Latency.zm = 10^(Multilink$Latency.approx.log)


#####################################################################
#####################################################################

######## Descriptive Statistics ########
########################################

## global ##
descstats1 = describe(Multilink)
descstats1 = as.data.frame(descstats1)

write.table(descstats1, file="descriptives.csv", append = FALSE, sep=",")

## direction ##
descstats2 = rbindlist(describeBy(Multilink$Latency, group=(Multilink$Direction)), use.names=TRUE,
fill=FALSE, idcol=TRUE)
descstats2 = rbind(descstats2,
  rbindlist(describeBy(Multilink$cycles, group=(Multilink$Direction)), use.names=TRUE, fill=FALSE,
idcol=TRUE),
  rbindlist(describeBy(Multilink$Stim.length, group=(Multilink$Direction)), use.names=TRUE,
fill=FALSE, idcol=TRUE),
  rbindlist(describeBy(Multilink$freq, group=(Multilink$Direction)), use.names=TRUE, fill=FALSE,
idcol=TRUE),
  rbindlist(describeBy(Multilink$Mean.N.RT, group=(Multilink$Direction)), use.names=TRUE,
fill=FALSE, idcol=TRUE),
  rbindlist(describeBy(Multilink$concreteness, group=(Multilink$Direction)), use.names=TRUE,
fill=FALSE, idcol=TRUE),
  rbindlist(describeBy(Multilink$Levdist, group=(Multilink$Direction)), use.names=TRUE, fill=FALSE,
idcol=TRUE),
  rbindlist(describeBy(Multilink$P.Latency.lm, group=(Multilink$Direction)), use.names=TRUE,
fill=FALSE, idcol=TRUE),
  rbindlist(describeBy(Multilink$P.Latency.zm, group=(Multilink$Direction)), use.names=TRUE,
fill=FALSE, idcol=TRUE) )
```

```
write.table(descstats2, file="descriptives.csv", append = TRUE, sep=",")


## the 8 conditions ##
descstats3 = rbindlist(describeBy(Multilink$Latency, group=(Multilink$category)))
descstats3 = rbind(descstats3,
  rbindlist(describeBy(Multilink$cycles, group=(Multilink$category))),
  rbindlist(describeBy(Multilink$Stim.length, group=(Multilink$category))),
  rbindlist(describeBy(Multilink$freq, group=(Multilink$category))),
  rbindlist(describeBy(Multilink$Mean.N.RT, group=(Multilink$category))),
  rbindlist(describeBy(Multilink$concreteness, group=(Multilink$category))),
  rbindlist(describeBy(Multilink$Levdist, group=(Multilink$category))),
  rbindlist(describeBy(Multilink$P.Latency.lm, group=(Multilink$category))),
  rbindlist(describeBy(Multilink$P.Latency.zm, group=(Multilink$category))) )


write.table(descstats3, file="descriptives.csv", append = TRUE, sep=",")

##################################################################
##################################################################


########## Correlational Analyses ##############
################################################

## data slices ##
# doesn't do correlations without defined slices, for some reason
Forward = Multilink$Direction=='Forward'
Backward = Multilink$Direction=='Backward'
ForwardHFCog = Multilink$category=='ForwardHFCog'
ForwardLFCog = Multilink$category=='ForwardLFCog'
ForwardHFNcog = Multilink$category=='ForwardHFNcog'
ForwardLFNcog = Multilink$category=='ForwardLFNcog'
BackwardHFCog = Multilink$category=='BackwardHFCog'
BackwardLFCog = Multilink$category=='BackwardLFCog'
BackwardHFNcog = Multilink$category=='BackwardHFNcog'
BackwardLFNcog = Multilink$category=='BackwardLFNcog'
```

```
## General correlations ##

## Latency & frequency ##
cor.latfreq = rbind.data.frame(
   tidy(cor.test(Multilink$Latency, Multilink$freq, use="complete.obs", method="spearman",
alternative= 'two.sided')),
   tidy(cor.test(Multilink$Latency[Forward], Multilink$freq[Forward], use='complete.obs',
method="spearman", alternative= 'two.sided')),
   tidy(cor.test(Multilink$Latency[Backward], Multilink$freq[Backward], use="complete.obs",
method="spearman", alternative= 'two.sided')),
   tidy(cor.test(Multilink$Latency[ForwardHFCog], Multilink$freq[ForwardHFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
   tidy(cor.test(Multilink$Latency[ForwardLFCog], Multilink$freq[ForwardLFCog], use="complete.obs",
method="spearman", alternative= 'two.sided')),
   tidy(cor.test(Multilink$Latency[ForwardHFNcog], Multilink$freq[ForwardHFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
   tidy(cor.test(Multilink$Latency[ForwardLFNcog], Multilink$freq[ForwardLFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
   tidy(cor.test(Multilink$Latency[BackwardHFCog], Multilink$freq[BackwardHFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
   tidy(cor.test(Multilink$Latency[BackwardLFCog], Multilink$freq[BackwardLFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
   tidy(cor.test(Multilink$Latency[BackwardHFNcog], Multilink$freq[BackwardHFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
   tidy(cor.test(Multilink$Latency[BackwardLFNcog], Multilink$freq[BackwardLFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')) )

rownames(cor.latfreq) <- paste(c("Global", "Forward", "Backward", "ForwardHFCog",
"ForwardLFCog", "ForwardHFNcog", "ForwardLFNcog", "BackwardHFCog", "BackwardLFCog",
"BackwardHFNcog", "BackwardLFNcog"), sep="")
write.table(cor.latfreq, file="correlations.csv", append = FALSE, sep=",")

## Cycles & Latency ##
cor.cyclat = rbind.data.frame(
  tidy(cor.test(Multilink$cycles, Multilink$Latency, use="complete.obs", method="spearman",
alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[Forward], Multilink$Latency[Forward], use="complete.obs",
method="spearman", alternative= 'two.sided')),
```

```
  tidy(cor.test(Multilink$cycles[Backward], Multilink$Latency[Backward], use="complete.obs",
method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[ForwardHFCog], Multilink$Latency[ForwardHFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[ForwardLFCog], Multilink$Latency[ForwardLFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[ForwardHFNcog], Multilink$Latency[ForwardHFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[ForwardLFNcog], Multilink$Latency[ForwardLFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[BackwardHFCog], Multilink$Latency[BackwardHFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[BackwardLFCog], Multilink$Latency[BackwardLFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[BackwardHFNcog], Multilink$Latency[BackwardHFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[BackwardLFNcog], Multilink$Latency[BackwardLFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')) )

rownames(cor.cyclat) <- paste(c("Global", "Forward", "Backward", "ForwardHFCog", "ForwardLFCog",
"ForwardHFNcog", "ForwardLFNcog", "BackwardHFCog", "BackwardLFCog", "BackwardHFNcog",
"BackwardLFNcog"), sep="")
write.table(cor.cyclat, file="correlations.csv", append = TRUE, sep=",")

## Cycles & frequency ##
cor.cycfreq = rbind.data.frame(
  tidy(cor.test(Multilink$cycles, Multilink$freq, use="complete.obs", method="spearman", alternative=
'two.sided')),
  tidy(cor.test(Multilink$cycles[Forward], Multilink$freq[Forward], use='complete.obs',
method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[Backward], Multilink$freq[Backward], use="complete.obs",
method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[ForwardHFCog], Multilink$freq[ForwardHFCog], use="complete.obs",
method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[ForwardLFCog], Multilink$freq[ForwardLFCog], use="complete.obs",
method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[ForwardHFNcog], Multilink$freq[ForwardHFNcog], use="complete.obs",
method="spearman", alternative= 'two.sided')),
```

```
  tidy(cor.test(Multilink$cycles[ForwardLFNcog], Multilink$freq[ForwardLFNcog], use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[BackwardHFCog], Multilink$freq[BackwardHFCog], use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[BackwardLFCog], Multilink$freq[BackwardLFCog], use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[BackwardHFNcog], Multilink$freq[BackwardHFNcog], use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[BackwardLFNcog], Multilink$freq[BackwardLFNcog], use="complete.obs", method="spearman", alternative= 'two.sided')) )

rownames(cor.cycfreq) <- paste(c("Global", "Forward", "Backward", "ForwardHFCog", "ForwardLFCog", "ForwardHFNcog", "ForwardLFNcog", "BackwardHFCog", "BackwardLFCog", "BackwardHFNcog", "BackwardLFNcog"), sep="")
write.table(cor.cycfreq, file="correlations.csv", append = TRUE, sep=",")

## Latency & Levenshtein Distance ##
cor.latlevd = rbind.data.frame(
  tidy(cor.test(Multilink$Latency, Multilink$Levdist, use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Latency[Forward], Multilink$Levdist[Forward], use='complete.obs', method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Latency[Backward], Multilink$Levdist[Backward], use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Latency[ForwardHFCog], Multilink$Levdist[ForwardHFCog], use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Latency[ForwardLFCog], Multilink$Levdist[ForwardLFCog], use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Latency[ForwardHFNcog], Multilink$Levdist[ForwardHFNcog], use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Latency[ForwardLFNcog], Multilink$Levdist[ForwardLFNcog], use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Latency[BackwardHFCog], Multilink$Levdist[BackwardHFCog], use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Latency[BackwardLFCog], Multilink$Levdist[BackwardLFCog], use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Latency[BackwardHFNcog], Multilink$Levdist[BackwardHFNcog], use="complete.obs", method="spearman", alternative= 'two.sided')),
```

```
  tidy(cor.test(Multilink$Latency[BackwardLFNcog], Multilink$Levdist[BackwardLFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')) )

rownames(cor.latlevd) <- paste(c("Global", "Forward", "Backward", "ForwardHFCog",
"ForwardLFCog", "ForwardHFNcog", "ForwardLFNcog", "BackwardHFCog", "BackwardLFCog",
"BackwardHFNcog", "BackwardLFNcog"), sep="")
write.table(cor.latlevd, file="correlations.csv", append = TRUE, sep=",")

## Cycles & Levenshtein Distance ##
cor.cyclevd = rbind.data.frame(
  tidy(cor.test(Multilink$cycles, Multilink$Levdist, use="complete.obs", method="spearman",
alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[Forward], Multilink$Levdist[Forward], use='complete.obs',
method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[Backward], Multilink$Levdist[Backward], use="complete.obs",
method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[ForwardHFCog], Multilink$Levdist[ForwardHFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[ForwardLFCog], Multilink$Levdist[ForwardLFCog], use="complete.obs",
method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[ForwardHFNcog], Multilink$Levdist[ForwardHFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[ForwardLFNcog], Multilink$Levdist[ForwardLFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[BackwardHFCog], Multilink$Levdist[BackwardHFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[BackwardLFCog], Multilink$Levdist[BackwardLFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[BackwardHFNcog], Multilink$Levdist[BackwardHFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[BackwardLFNcog], Multilink$Levdist[BackwardLFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')) )

rownames(cor.cyclevd) <- paste(c("Global", "Forward", "Backward", "ForwardHFCog",
"ForwardLFCog", "ForwardHFNcog", "ForwardLFNcog", "BackwardHFCog", "BackwardLFCog",
"BackwardHFNcog", "BackwardLFNcog"), sep="")
write.table(cor.cyclevd, file="correlations.csv", append = TRUE, sep=",")

## Frequency & Levenshtein Distance ##
```

```
cor.freqlevd = rbind.data.frame(
  tidy(cor.test(Multilink$freq, Multilink$Levdist, use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$freq[Forward], Multilink$Levdist[Forward], use='complete.obs', method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$freq[Backward], Multilink$Levdist[Backward], use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$freq[ForwardHFCog], Multilink$Levdist[ForwardHFCog], use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$freq[ForwardLFCog], Multilink$Levdist[ForwardLFCog], use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$freq[ForwardHFNcog], Multilink$Levdist[ForwardHFNcog], use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$freq[ForwardLFNcog], Multilink$Levdist[ForwardLFNcog], use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$freq[BackwardHFCog], Multilink$Levdist[BackwardHFCog], use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$freq[BackwardLFCog], Multilink$Levdist[BackwardLFCog], use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$freq[BackwardHFNcog], Multilink$Levdist[BackwardHFNcog], use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$freq[BackwardLFNcog], Multilink$Levdist[BackwardLFNcog], use="complete.obs", method="spearman", alternative= 'two.sided')) )

rownames(cor.freqlevd) <- paste(c("Global", "Forward", "Backward", "ForwardHFCog", "ForwardLFCog", "ForwardHFNcog", "ForwardLFNcog", "BackwardHFCog", "BackwardLFCog", "BackwardHFNcog", "BackwardLFNcog"), sep="")
write.table(cor.freqlevd, file="correlations.csv", append = TRUE, sep=",")


## Linear model correlations ##

## Cycles & P.Latency.lm ##
cor.cycLM = rbind.data.frame(
  tidy(cor.test(Multilink$cycles, Multilink$P.Latency.lm, use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[Forward], Multilink$P.Latency.lm[Forward], use="complete.obs", method="spearman", alternative= 'two.sided')),
```

```
  tidy(cor.test(Multilink$cycles[Backward], Multilink$P.Latency.lm[Backward], use="complete.obs",
method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[ForwardHFCog], Multilink$P.Latency.lm[ForwardHFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[ForwardLFCog], Multilink$P.Latency.lm[ForwardLFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[ForwardHFNcog], Multilink$P.Latency.lm[ForwardHFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[ForwardLFNcog], Multilink$P.Latency.lm[ForwardLFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[BackwardHFCog], Multilink$P.Latency.lm[BackwardHFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[BackwardLFCog], Multilink$P.Latency.lm[BackwardLFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[BackwardHFNcog], Multilink$P.Latency.lm[BackwardHFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$cycles[BackwardLFNcog], Multilink$P.Latency.lm[BackwardLFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')) )

rownames(cor.cycLM) <- paste(c("Global", "Forward", "Backward", "ForwardHFCog",
"ForwardLFCog", "ForwardHFNcog", "ForwardLFNcog", "BackwardHFCog", "BackwardLFCog",
"BackwardHFNcog", "BackwardLFNcog"), sep="")
write.table(cor.cycLM, file="correlations.csv", append = TRUE, sep=",")

## Frequency & P.Latency.lm ##
cor.freqLM = rbind.data.frame(
  tidy(cor.test(Multilink$freq, Multilink$P.Latency.lm, use="complete.obs", method="spearman",
alternative= 'two.sided')),
  tidy(cor.test(Multilink$freq[Forward], Multilink$P.Latency.lm[Forward], use='complete.obs',
method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$freq[Backward], Multilink$P.Latency.lm[Backward], use="complete.obs",
method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$freq[ForwardHFCog], Multilink$P.Latency.lm[ForwardHFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$freq[ForwardLFCog], Multilink$P.Latency.lm[ForwardLFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$freq[ForwardHFNcog], Multilink$P.Latency.lm[ForwardHFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
```

```
  tidy(cor.test(Multilink$freq[ForwardLFNcog], Multilink$P.Latency.lm[ForwardLFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$freq[BackwardHFCog], Multilink$P.Latency.lm[BackwardHFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$freq[BackwardLFCog], Multilink$P.Latency.lm[BackwardLFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$freq[BackwardHFNcog], Multilink$P.Latency.lm[BackwardHFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$freq[BackwardLFNcog], Multilink$P.Latency.lm[BackwardLFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')) )

rownames(cor.freqLM) <- paste(c("Global", "Forward", "Backward", "ForwardHFCog",
"ForwardLFCog", "ForwardHFNcog", "ForwardLFNcog", "BackwardHFCog", "BackwardLFCog",
"BackwardHFNcog", "BackwardLFNcog"), sep="")
write.table(cor.freqLM, file="correlations.csv", append = TRUE, sep=",")

## Latency & P.Latency.lm ##
cor.latLM = rbind.data.frame(
  tidy(cor.test(Multilink$Latency, Multilink$P.Latency.lm, use="complete.obs", method="spearman",
alternative= 'two.sided')),
  tidy(cor.test(Multilink$Latency[Forward], Multilink$P.Latency.lm[Forward], use='complete.obs',
method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Latency[Backward], Multilink$P.Latency.lm[Backward], use="complete.obs",
method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Latency[ForwardHFCog], Multilink$P.Latency.lm[ForwardHFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Latency[ForwardLFCog], Multilink$P.Latency.lm[ForwardLFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Latency[ForwardHFNcog], Multilink$P.Latency.lm[ForwardHFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Latency[ForwardLFNcog], Multilink$P.Latency.lm[ForwardLFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Latency[BackwardHFCog], Multilink$P.Latency.lm[BackwardHFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Latency[BackwardLFCog], Multilink$P.Latency.lm[BackwardLFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Latency[BackwardHFNcog], Multilink$P.Latency.lm[BackwardHFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
```

```
  tidy(cor.test(Multilink$Latency[BackwardLFNcog], Multilink$P.Latency.lm[BackwardLFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')) )

rownames(cor.latLM) <- paste(c("Global", "Forward", "Backward", "ForwardHFCog", "ForwardLFCog",
"ForwardHFNcog", "ForwardLFNcog", "BackwardHFCog", "BackwardLFCog", "BackwardHFNcog",
"BackwardLFNcog"), sep="")
write.table(cor.latLM, file="correlations.csv", append = TRUE, sep=",")

## Levenshtein Distance & P.Latency.lm ##
cor.levdLM = rbind.data.frame(
  tidy(cor.test(Multilink$Levdist, Multilink$P.Latency.lm, use="complete.obs", method="spearman",
alternative= 'two.sided')),
  tidy(cor.test(Multilink$Levdist[Forward], Multilink$P.Latency.lm[Forward], use='complete.obs',
method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Levdist[Backward], Multilink$P.Latency.lm[Backward], use="complete.obs",
method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Levdist[ForwardHFCog], Multilink$P.Latency.lm[ForwardHFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Levdist[ForwardLFCog], Multilink$P.Latency.lm[ForwardLFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Levdist[ForwardHFNcog], Multilink$P.Latency.lm[ForwardHFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Levdist[ForwardLFNcog], Multilink$P.Latency.lm[ForwardLFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Levdist[BackwardHFCog], Multilink$P.Latency.lm[BackwardHFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Levdist[BackwardLFCog], Multilink$P.Latency.lm[BackwardLFCog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Levdist[BackwardHFNcog], Multilink$P.Latency.lm[BackwardHFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')),
  tidy(cor.test(Multilink$Levdist[BackwardLFNcog], Multilink$P.Latency.lm[BackwardLFNcog],
use="complete.obs", method="spearman", alternative= 'two.sided')) )

rownames(cor.levdLM) <- paste(c("Global", "Forward", "Backward", "ForwardHFCog",
"ForwardLFCog", "ForwardHFNcog", "ForwardLFNcog", "BackwardHFCog", "BackwardLFCog",
"BackwardHFNcog", "BackwardLFNcog"), sep="")
write.table(cor.levdLM, file="correlations.csv", append = TRUE, sep=",")
```

```
#########################################################################
#########################################################################

####################### Correlation Graphs ##############################
###########################################################################
# AKA Corrgrams (but not correlograms, thats actually different)
# uses the 'corrgram' package


## handy-dandy script below ##
# makes the corrgrams look nicer
# found on : http://stackoverflow.com/questions/19012529/correlation-corrplot-configuration

###########################################

panel.shadeNtext <- function (x, y, corr = NULL, col.regions, ...)
{
  results <- cor.test(x, y, use="complete.obs", alternative = "two.sided", method='spearman')
  est <- results$p.value
  stat <- results$estimate
  stars <- ifelse(est < 1e-4, "p = < 0.0001",
                  ifelse(est < 1e-3, "p = 0.001",
                         ifelse(est < 1e-2, "p = 0.01",
                                ifelse(est < 5e-2, "p = 0.05",
                                       ifelse(est < 1e-1, "p = 0.1", "NS")))))
  ncol <- 14
  pal <- col.regions(ncol)
  col.ind <- as.numeric(cut(stat, breaks = seq(from = -1, to = 1,
                                               length = ncol + 1), include.lowest = TRUE))
  usr <- par("usr")
  rect(usr[1], usr[3], usr[2], usr[4], col = pal[col.ind],
       border = NA)
  box(col = "lightgray")
  on.exit(par(usr))
  par(usr = c(0, 1, 0, 1))
  r <- formatC(stat, digits= 2, format= 'f')
  cex.cor <- .8/strwidth("-X.xx")
  cex.star <- .5/strwidth("p = x.xxx")
```

```
  fonts <- ifelse(stars != "", 2,1)
  # option 1: stars:
  text(0.5, 0.7, paste0(r), cex = cex.cor)
  text(0.5, 0.2, paste0(stars), cex = cex.star)
  # option 2: bolding:
  #text(0.5, 0.5, r, cex = cex.cor, font=fonts)
}

############################################################

# Call the corrgram function with the new panel functions
# NB: call on the data, not the correlation matrix

## new data frame ##
# main variables of numeric-type only, so placed into a new data frame.

numdframe = cbind.data.frame(Multilink$Latency, Multilink$cycles, Multilink$freq,
Multilink$P.Latency.lm, Multilink$P.Latency.zm, Multilink$Levdist, Multilink$Direction,
Multilink$category)
colnames(numdframe) <- paste(c("Latency", "Cycle-time", "Frequency", "L-scaled", "Z-scaled", "LD",
"Direction", "Category"), sep="")

# best colour choices as determined by Jakob, Thomas, Sean, & myself:
# col.regions=colorRampPalette(c('blue', 'white', 'red'))   (visually-informative)
# col.regions=colorRampPalette(c('blue', 'red'))            (looks cool?)
# col.regions=colorRampPalette(c('blue', 'green'))          (my favourite)


## global ##
png(filename= 'Figure xx Corregram Global.png', width= 1200, height= 1066)
corrgram(numdframe, type='data', lower.panel=panel.shadeNtext,
      upper.panel=NULL, cex.labels= 5.0, col.regions=colorRampPalette(c('blue', 'white', 'red')),
cor.method='spearman')
title('Global Condition', cex.main=4.5)
dev.off()

## Forward ##
png(filename= 'Figure xx Corregram Forward.png', width= 1200, height= 1066)
```

```
corrgram(numdframe[numdframe$Direction=='Forward',], type='data', lower.panel=panel.shadeNtext,
        upper.panel=NULL, cex.labels= 5.0, col.regions=colorRampPalette(c('blue', 'white', 'red')),
cor.method='spearman')
title('Forward Condition', cex.main=4.5)
dev.off()

## Backward ##
png(filename= 'Figure xx Corregram Backward.png', width= 1200, height= 1066)
corrgram(numdframe[numdframe$Direction=='Backward',], type='data',
lower.panel=panel.shadeNtext,
        upper.panel=NULL, cex.labels= 5.0, col.regions=colorRampPalette(c('blue', 'white', 'red')),
cor.method='spearman')
title('Backward Condition', cex.main=4.5)
dev.off()


# without the predicted-latencies (since correlations are the same)
numdframe2 = cbind.data.frame(Multilink$Latency, Multilink$cycles, Multilink$freq, Multilink$Levdist,
Multilink$Direction, Multilink$category)
colnames(numdframe2) <- paste(c("Latency", "Cycle-time", "Frequency", "LD", "Direction",
"Category"), sep="")

## Forward HF Cognate ##
png(filename= 'Figure xx Corregram ForwardHFCog.png', width= 1200, height= 1066)
corrgram(numdframe2[numdframe2$Category=='ForwardHFCog',], type='data',
lower.panel=panel.shadeNtext,
        upper.panel=NULL, cex.labels= 7.0, col.regions=colorRampPalette(c('blue', 'white', 'red')),
cor.method='spearman')
title('Forward High-Frequency Cognate Condition', cex.main=4.0)
dev.off()

## Forward LF Cognate ##
png(filename= 'Figure xx Corregram ForwardLFCog.png', width= 1200, height= 1066)
corrgram(numdframe2[numdframe2$Category=='ForwardLFCog',], type='data',
lower.panel=panel.shadeNtext,
        upper.panel=NULL, cex.labels= 7.0, col.regions=colorRampPalette(c('blue', 'white', 'red')),
cor.method='spearman')
title('Forward Low-Frequency Cognate Condition', cex.main=4.0)
dev.off()
```

```
## Forward HF Noncognate ##
png(filename= 'Figure xx Corregram ForwardHFNcog.png', width= 1200, height= 1066)
corrgram(numdframe2[numdframe2$Category=='ForwardHFNcog',], type='data',
lower.panel=panel.shadeNtext,
        upper.panel=NULL, cex.labels= 7.0, col.regions=colorRampPalette(c('blue', 'white', 'red')),
cor.method='spearman')
title('Forward High-Frequency Noncognate Condition', cex.main=4.0)
dev.off()

## Forward LF Noncognate ##
png(filename= 'Figure xx Corregram ForwardLFNcog.png', width= 1200, height= 1066)
corrgram(numdframe2[numdframe2$Category=='ForwardLFNcog',], type='data',
lower.panel=panel.shadeNtext,
        upper.panel=NULL, cex.labels= 7.0, col.regions=colorRampPalette(c('blue', 'white', 'red')),
cor.method='spearman')
title('Forward Low-Frequency Noncognate Condition', cex.main=4.0)
dev.off()


## Backward HF Cognate ##
png(filename= 'Figure xx Corregram BackwardHFCog.png', width= 1200, height= 1066)
corrgram(numdframe2[numdframe2$Category=='BackwardHFCog',], type='data',
lower.panel=panel.shadeNtext,
        upper.panel=NULL, cex.labels= 7.0, col.regions=colorRampPalette(c('blue', 'white', 'red')),
cor.method='spearman')
title('Backward High-Frequency Cognate Condition', cex.main=4.0)
dev.off()

## Backward LF Cognate ##
png(filename= 'Figure xx Corregram BackwardLFCog.png', width= 1200, height= 1066)
corrgram(numdframe2[numdframe2$Category=='BackwardLFCog',], type='data',
lower.panel=panel.shadeNtext,
        upper.panel=NULL, cex.labels= 7.0, col.regions=colorRampPalette(c('blue', 'white', 'red')),
cor.method='spearman')
title('Backward Low-Frequency Cognate Condition', cex.main=4.0)
dev.off()

## Backward HF Noncognate ##
```

```
png(filename= 'Figure xx Corregram BackwardHFNcog.png', width= 1200, height= 1066)
corrgram(numdframe2[numdframe2$Category=='BackwardHFNcog',], type='data',
lower.panel=panel.shadeNtext,
        upper.panel=NULL, cex.labels= 7.0, col.regions=colorRampPalette(c('blue', 'white', 'red')),
cor.method='spearman')
title('Backward High-Frequency Noncognate Condition', cex.main=4.0)
dev.off()

## Backward LF Noncognate ##
png(filename= 'Figure xx Corregram BackwardLFNcog.png', width= 1200, height= 1066)
corrgram(numdframe2[numdframe2$Category=='BackwardLFNcog',], type='data',
lower.panel=panel.shadeNtext,
        upper.panel=NULL, cex.labels= 7.0, col.regions=colorRampPalette(c('blue', 'white', 'red')),
cor.method='spearman')
title('Backward Low-Frequency Noncognate Condition', cex.main=4.0)
dev.off()




## Means correlations ##
# intermediate condition examining between-group correlations
# standard way of doing things, I guess.

newdf <- ddply(Multilink, c('Direction', 'cognacy', 'freq.cat', 'category'), summarize,
Latmean=mean(Latency), cycmean=mean(cycles), freqmean=mean(freq), ldmean=mean(Levdist))

cor.mean <- rbind.data.frame(
  tidy(cor.test(newdf$Latmean, newdf$cycmean, use="complete.obs", method="spearman",
alternative= 'two.sided')),
  tidy(cor.test(newdf$Latmean, newdf$freqmean, use="complete.obs", method="spearman",
alternative= 'two.sided')),
  tidy(cor.test(newdf$freqmean, newdf$cycmean, use="complete.obs", method="spearman",
alternative= 'two.sided')),
  tidy(cor.test(newdf$ldmean, newdf$cycmean, use="complete.obs", method="spearman",
alternative= 'two.sided')),
  tidy(cor.test(newdf$freqmean, newdf$ldmean, use="complete.obs", method="spearman",
alternative= 'two.sided')),
  tidy(cor.test(newdf$Latmean, newdf$ldmean, use="complete.obs", method="spearman", alternative=
'two.sided')))
```

```
rownames(cor.mean) <- c('Latency * Cycle-time', 'Latency * Frequency', 'Frequency * Cycle-time', 'LD
* Cycle-time', 'Frequency * LD', 'Latency * LD')
write.table(cor.mean, file="correlations.csv", append = TRUE, sep=",")


colnames(newdf)[5:8] <- c('Latency', 'Cycle-time', 'Frequency', 'LD')

png(filename= 'Figure xx corregram mean-regressed.png', width= 1200, height= 1066)
corrgram(newdf, type='data', lower.panel=panel.shadeNtext,
        upper.panel=NULL, cex.labels= 7.0, col.regions=colorRampPalette(c('blue', 'white', 'red')),
cor.method='spearman')
title('Mean-regressed Correlations', cex.main=4.0)
dev.off()


#######################################################

############### Analysis Of Variance #################
#######################################################
# Partial Omega Squared obtained from script:
# http://pastebin.com/raw.php?i=iA6CqQF9
# included up top, just highlight & source\run
# 'lsr' is included as a library if one wishes to calculate eta ES
# other possible ES measurements available in 'EffectSizeStats.R'

## Latency ##
Latencyanova <- aov(Latency ~ Direction*cognacy*freq.cat, data= Multilink)
summary(Latencyanova)
Omega1 <- partialOmegas(Latencyanova)

## Cycle-time ##
Cycleanova <- aov(cycles ~ Direction*cognacy*freq.cat, data= Multilink)
summary(Cycleanova)
Omega2 <-partialOmegas(Cycleanova)

## P.Latency.lm ##
LManova <- aov(P.Latency.lm ~ Direction*cognacy*freq.cat, data= Multilink)
```

```
summary(LManova)
Omega3 <-partialOmegas(LManova)

## P.Latency.zm ##
ZManova <- aov(P.Latency.zm ~ Direction*cognacy*freq.cat, data= Multilink)
summary(ZManova)
Omega4 <- partialOmegas(ZManova)

Allomegas <- data.table(Omega1, Omega2, Omega3, Omega4)
rownames(Allomegas) <-  paste(c('direction', 'cognacy', 'freq.cat', 'direction*cognacy',
'direction*freq.cat', 'cognacy*freq.cat', 'direction*cognacy*freq.cat'))
colnames(Allomegas) <-  paste(c('Latency', 'Cycle-time', 'L-scaled', 'Z-scaled'))

anovadf <- rbind.data.frame(
  tidy(Latencyanova),
  tidy(Cycleanova),
  tidy(LManova),
  tidy(ZManova))

write.table(anovadf, file='ANOVA.csv', append=FALSE, sep=',', na='', row.names=TRUE,
col.names=TRUE)
write.table(Allomegas, file='PartialOmegas.csv', append=FALSE, sep=',', na='', row.names=TRUE,
col.names=TRUE)


##########################################################

############### Generalized Linear Regression (analyzing categorical effects) ###############
##########################################################################################
############
#"intercepts" for primary tests are the Forward\Cognate\High-frequency combinations#

# unused
Lowfunction1 <- cbind.data.frame(Multilink$Latency, Multilink$Direction, Multilink$cognacy,
Multilink$freq.cat, Multilink$Ph.Onset, Multilink$Stim.length, Multilink$concreteness)
Lowfunction1 <- data.matrix(Lowfunction1)
colnames(Lowfunction1) <-  paste(c('Latency', 'Direction', 'cognacy', 'freq.cat', 'Ph.Onset',
'Stim.length', 'concreteness'))
```

```
Lowfunction1 <- as.data.frame(Lowfunction1)
##
#

# model selection & comparison
# GAM model, #3, has the lowest AIC.
model1 <- glm(Latency ~ (Direction+cognacy+freq.cat+Ph.Onset+Stim.length+concreteness)^2,
data= Multilink)
model2 <- lmer(Latency~
(Direction+cognacy+freq.cat+(1|Ph.Onset)+(1|Stim.length)+(1|concreteness))^2, REML=TRUE, data=
Multilink)
model3 <- gam(Latency~ (Direction+cognacy+freq.cat+Ph.Onset+Stim.length+concreteness)^2,
data= Multilink, family= gaussian, method='REML', select=TRUE, fit=TRUE, scale=8000)
model4 <- gamlss(Latency~ (Direction+cognacy+freq.cat+Ph.Onset+Stim.length+concreteness)^2,
data= Multilink, family= NO(), method=RS())
model5 <- gamm4(Latency~ (Direction+cognacy+freq.cat)^2, random=~
(1|Ph.Onset)+(1|Stim.length)+(1|concreteness), REML=TRUE, data= Multilink, family= gaussian())

summary(model1)
summary(model2)
summary(model3)
summary(model4)
summary(model5$mer)


## Latency ##
Latencygam <- gam(Latency~ (Direction+cognacy+freq.cat+Ph.Onset+Stim.length+concreteness)^2,
data= Multilink, family= gaussian, method='GCV.Cp', select=TRUE, fit=TRUE, scale=8500)
summary.latgam0 <- summary.gam(Latencygam)
print(summary.latgam0)
Latencygam.coef <- as.data.frame(summary.latgam0$p.table)
Latencygam.chisq <- as.data.frame(summary.latgam0$pTerms.table)
Latencygam.aov <- as.data.frame(tidy(aov(Latencygam)), row.names=NULL)

write.table(Latencygam.coef, file='GAM.csv', append=FALSE, sep=',', na='', row.names=TRUE,
col.names=TRUE)
write.table(Latencygam.chisq, file='GAM.csv', append=TRUE, sep=',', na='', row.names=TRUE,
col.names=TRUE)
```

```
write.table(Latencygam.aov, file='GAM.csv', append=TRUE, sep=',', na='', row.names=TRUE,
col.names=TRUE)


## Cycle-time ##
Cyclegam <- gam(cycles~ (Direction+cognacy+freq.cat+Ph.Onset+Stim.length+concreteness)^2,
data= Multilink, family= gaussian, method='GCV.Cp', select=TRUE, fit=TRUE, scale=2)
summary.cycgam0 <- summary.gam(Cyclegam)
print(summary.cycgam0)
cyclegam.coef <- as.data.frame(summary.cycgam0$p.table)
cyclegam.chisq <- as.data.frame(summary.cycgam0$pTerms.table)
cyclegam.aov <- as.data.frame(tidy(aov(Cyclegam)), row.names=NULL)

write.table(cyclegam.coef, file='GAM.csv', append=TRUE, sep=',', na='', row.names=TRUE,
col.names=TRUE)
write.table(cyclegam.chisq, file='GAM.csv', append=TRUE, sep=',', na='', row.names=TRUE,
col.names=TRUE)
write.table(cyclegam.aov, file='GAM.csv', append=TRUE, sep=',', na='', row.names=TRUE,
col.names=TRUE)


## P.Latency.lm ##
LMgam <- gam(P.Latency.lm~ (Direction+cognacy+freq.cat+Ph.Onset+Stim.length+concreteness)^2,
data= Multilink, family= gaussian, method='GCV.Cp', select=TRUE, fit=TRUE, scale=1400)
summary.LMgam0 <- summary.gam(LMgam)
print(summary.LMgam0)
LMgam.coef <- as.data.frame(summary.LMgam0$p.table)
LMgam.chisq <- as.data.frame(summary.LMgam0$pTerms.table)
LMgam.aov <- as.data.frame(tidy(aov(LMgam)), row.names=NULL)

write.table(LMgam.coef, file='GAM.csv', append=TRUE, sep=',', na='', row.names=TRUE,
col.names=TRUE)
write.table(LMgam.chisq, file='GAM.csv', append=TRUE, sep=',', na='', row.names=TRUE,
col.names=TRUE)
write.table(LMgam.aov, file='GAM.csv', append=TRUE, sep=',', na='', row.names=TRUE,
col.names=TRUE)


## P.Latency.zm ##
```

```
ZMgam <- gam(P.Latency.zm~ (Direction+cognacy+freq.cat+Ph.Onset+Stim.length+concreteness)^2,
data= Multilink, family= gaussian, method='GCV.Cp', select=TRUE, fit=TRUE, scale=7500)
summary.ZMgam0 <- summary.gam(ZMgam)
print(summary.ZMgam0)
ZMgam.coef <- as.data.frame(summary.ZMgam0$p.table)
ZMgam.chisq <- as.data.frame(summary.ZMgam0$pTerms.table)
ZMgam.aov <- as.data.frame(tidy(aov(ZMgam)), row.names=NULL)

write.table(ZMgam.coef, file='GAM.csv', append=TRUE, sep=',', na='', row.names=TRUE,
col.names=TRUE)
write.table(ZMgam.chisq, file='GAM.csv', append=TRUE, sep=',', na='', row.names=TRUE,
col.names=TRUE)
write.table(ZMgam.aov, file='GAM.csv', append=TRUE, sep=',', na='', row.names=TRUE,
col.names=TRUE)


##########################################################

########### Model-Approximate X2 Analysis ############
#####################################################################################
####
# assesses the fitness of model-data to empirical-data
# primary test is the Model-approximate X2, essentially a statistical-distance metric

## Linear Model assessment ##

#stored variables: Data & Linear Model Latency and sample variance
Latencydata = Multilink$Latency
LatencymodelLM = Multilink$P.Latency.lm
samplevardata = var(Multilink$Latency)
samplevarmodelLM = var(Multilink$P.Latency.lm)

#Model-Approximate X2, Linear Model
X2sumLM = (sum((Latencydata - LatencymodelLM)^2) / (((samplevardata^2) / nrow(Multilink)) +
((samplevarmodelLM^2) / nrow(Multilink))))
print(X2sumLM)
# Degrees of Freedom = 247 (N [sample size] - p [free parameters in the model])
# 249 paired datapoints, - 2 parameters [latency, cycle-time] = DF 247
```

```
# get the P-value from an online X2 calculator, should be around 0.9
# but that's good, showing that the divergence
# Highlight + Ctrl-Enter Lines XXX-XXX (Linear Model Assessment to above line) to run properly
# go here for a chi-square p-value calculator: http://stattrek.com/online-calculator/chi-square.aspx

## Kolmogorov-Smirnov test ##
#by running both 'less', and 'greater', you can better gain a sense of where the estimate is
#x is the prediction, y is the match
ks.test(Multilink$P.Latency.lm, Multilink$Latency, alternative = 'two.sided')
ks.test(Multilink$P.Latency.lm, Multilink$Latency, alternative = 'less')
ks.test(Multilink$P.Latency.lm, Multilink$Latency, alternative = 'greater')

## Kullback-Leibler Divergence ##
# measures goodness-of-fit through information loss
# asymmetric test, must be run in both directions then averaged
#'log2' produces the measurement in Shannons\bits
# Uses the 'seewave' package
kl.dist(Multilink$Latency, Multilink$P.Latency.lm, base =2)
kl.dist(Multilink$cycles, Multilink$P.Latency.lm, base =2)

## Empirical X2 ##
# goes with model-approximate X2 test, backs up the results
# compares sample distributions on a X2 distribution
# output in information-theoretic relevant quantities
# allows comparison to the other 2 statistics.
chi2.empirical(Multilink$Latency, Multilink$P.Latency.lm, unit = 'log2')
chi2.empirical(Multilink$Latency, Multilink$cycles, unit = 'log2')
chi2.empirical(Multilink$cycles, Multilink$P.Latency.lm, unit = 'log2')

## Two-sample Density-based Empirical Likelihood ##
# unused in paper, but interesting
# uses the 'dbEMPLikeGOF' package
# basically measures uniformity and distribution equality, forming a distance measure
dbEmpLikeGOF(Multilink$Latency, Multilink$P.Latency.lm)
dbEmpLikeGOF(Multilink$Latency, Multilink$cycles)
dbEmpLikeGOF(Multilink$cycles, Multilink$P.Latency.lm)

## Fan Test & Geometric test ##
```

```
# another measure of statistical distance
# first value is the model distribution (null hypothesis)
# second value is the empirical distribution
# output estimates the total distance through density-estimations
# Null-h = CDF 1 does have the same density\distance as CDF 2
# doesn't seem to be too useful, unused, whatever.
fan.test(Multilink$P.Latency.lm, ecdf(Multilink$Latency))
fan.test(Multilink$cycles, ecdf(Multilink$Latency))
fan.test(Multilink$cycles, ecdf(Multilink$P.Latency.lm))

dgeometric.test(Multilink$P.Latency.lm, ecdf(Multilink$Latency))
dgeometric.test(Multilink$cycles, ecdf(Multilink$Latency))
dgeometric.test(Multilink$cycles, ecdf(Multilink$P.Latency.lm))


## Z-score Model assessment ##

#stored variables: Data & Linear Model Latency and sample variance
Latencydata = Multilink$Latency
LatencymodelZM = Multilink$P.Latency.zm
samplevardata = var(Multilink$Latency)
samplevarmodelZM = var(Multilink$P.Latency.zm)

#Model-Approximate X2, Z-score Model
X2sumZM = (sum((Latencydata - LatencymodelZM)^2) / (((samplevardata^2) / nrow(Multilink)) +
((samplevarmodelZM^2) / nrow(Multilink))))
print(X2sumZM)
# Degrees of Freedom = 247 (N [sample size] - p [free parameters in the model])
# 249 paired datapoints, - 2 parameters [latency, cycle-time] = DF 247
# get the P-value from an online X2 calculator, should be around 0.9
# but that's good, showing that the divergence
# Highlight + Ctrl-Enter Lines XXX-XXX (Linear Model Assessment to above line) to run properly

## Kolmogorov-Smirnov test ##
#by running both 'less', and 'greater', you can better gain a sense of where the estimate is
#x is the prediction, y is the match
ks.test(Multilink$P.Latency.zm, Multilink$Latency, alternative = 'two.sided')
ks.test(Multilink$P.Latency.zm, Multilink$Latency, alternative = 'less')
```

```
ks.test(Multilink$P.Latency.zm, Multilink$Latency, alternative = 'greater')

## Kullback-Leibler Divergence ##
# measures goodness-of-fit through information loss
#'log2' produces the measurement in Shannons\'bits'
# Uses the 'seewave' package
kl.dist(Multilink$Latency, Multilink$P.Latency.zm, base=2)
kl.dist(Multilink$Latency, Multilink$cycles, base=2)
kl.dist(Multilink$cycles, Multilink$P.Latency.zm, base=2)

## Empirical X2 ##
# unused in paper. goes with KL-D test, backs up the results
# compares sample distributions on a X2 distribution, and measures the divergence from both
samples
# pretty much comparable to the Model-approximate X2, but measures in Shannons
chi2.empirical(Multilink$Latency, Multilink$P.Latency.zm, unit = 'log2')
chi2.empirical(Multilink$Latency, Multilink$cycles, unit = 'log2')
chi2.empirical(Multilink$cycles, Multilink$P.Latency.zm, unit = 'log2')

## Two-sample Density-based Empirical Likelihood ##
# unused in paper, but interesting
# uses the {dbEMPLikeGOF} package
# basically measures uniformity and distribution equality, forming a distance measure
dbEmpLikeGOF(Multilink$Latency, Multilink$P.Latency.zm)
dbEmpLikeGOF(Multilink$Latency, Multilink$cycles)
dbEmpLikeGOF(Multilink$cycles, Multilink$P.Latency.zm)

## Fan Test & Geometric test ##
# another measure of statistical distance
# first value is the model distribution (null hypothesis)
# second value is the empirical distribution
# output estimates the total distance through density-estimations
# Null-h = CDF 1 does have the same density\distance as CDF 2
# unused, doesn't seem to be too useful, whatever.
fan.test(Multilink$P.Latency.zm, ecdf(Multilink$Latency))
fan.test(Multilink$cycles, ecdf(Multilink$Latency))
fan.test(Multilink$cycles, ecdf(Multilink$P.Latency.zm))
```

```
dgeometric.test(Multilink$P.Latency.zm, ecdf(Multilink$Latency))
dgeometric.test(Multilink$cycles, ecdf(Multilink$Latency))
dgeometric.test(Multilink$cycles, ecdf(Multilink$P.Latency.zm))


#####################################################

########### Shaprio-Wilks Normality ############
################################################

## High frequency ##
shapiro.test(Multilink$freq[Multilink$freq.cat=="High-freq"])

## Low frequency ##
shapiro.test(Multilink$freq[Multilink$freq.cat=="Low-freq"])

## All frequency ##
shapiro.test(Multilink$freq)

## Latency ##
shapiro.test(Multilink$Latency)

## Cycle-time ##
shapiro.test(Multilink$cycles)

## P.Latency.lm ##
shapiro.test(Multilink$P.Latency.lm)

## P.Latency.zm ##
shapiro.test(Multilink$P.Latency.zm)


###################################################

########      Graphs       ########
####################################
#par(mfrow = c(2,1))
```

```
## cumulative distribution function graph ##
# compares model and empirical CDFs
# highlight each section to run

# Latency & LM
ecdf.l.lmapprox = ecdf(Multilink$P.Latency.lm)
ecdf.l = ecdf(Multilink$Latency)
plot(ecdf.l, ylab='Datapoint Cumulative Distribution', xlab='Milliseconds', main='Latency * Linear Model CDF')
xx = knots(ecdf.l.lmapprox)
xx = 600:1200
lines(xx,ecdf.l.lmapprox(xx), col=2, lty=2)


# Latency & ZM
ecdf.l.zmapprox = ecdf(Multilink$P.Latency.zm)
ecdf.l = ecdf(Multilink$Latency)
plot(ecdf.l, ylab='Datapoint Cumulative Distribution', xlab='Milliseconds', main='Latency * Z-score Model CDF')
xx = knots(ecdf.l.zmapprox)
xx = 600:1200
lines(xx,ecdf.l.zmapprox(xx), col=2, lty=2)


## Linear model scatterplot ##
# uses package 'ggplot2'
# basically the same thing as the regression plots below
# highlight each section to run

ggplot(Multilink, aes(x = Latency, y = cycles)) +
  labs(y= 'cycle-time', title= 'Linear Relationship') +
  geom_point() +
  stat_smooth(method = "lm", se=TRUE, col = "red")

png(filename='Figure xx Linear model scaling.png', width=1200, height=1066)
ggplot(Multilink, aes(x = Latency, y = P.Latency.lm)) +
  labs(y= 'L-scaled Latency', title= 'Linear Model Scaling') +
  geom_point() +
```

```
  stat_smooth(method = "lm", se=TRUE, col = "red")+
  theme(plot.title = element_text(face="bold", size=28),
        axis.text.x= element_text(size=20),
        axis.text.y= element_text(size=20),
        legend.title= element_text(size=22),
        legend.text= element_text(size=18),
        axis.title.x = element_text(size=26),
        axis.title.y = element_text(size=26))
dev.off()

png(filename='Figure xx Z-score model scaling.png', width=1200, height=1066)
ggplot(Multilink, aes(x = Latency, y = P.Latency.zm)) +
  labs(y= 'Z-scaled Latency', title= 'Z-score Model Scaling') +
  geom_point() +
  stat_smooth(method = "lm", se=TRUE, col = "red")+
  theme(plot.title = element_text(face="bold", size=28),
        axis.text.x= element_text(size=20),
        axis.text.y= element_text(size=20),
        legend.title= element_text(size=22),
        legend.text= element_text(size=18),
        axis.title.x = element_text(size=26),
        axis.title.y = element_text(size=26))
dev.off()

## Regression model graphs ##
# basically the same as the ANOVA graphs

lateffect.cogstim <- ddply(Multilink, c('cognacy', 'Stim.length'), summarize,
AVERAGE=mean(Latency), SE=sqrt(var(Latency)/length(Latency)))
lateffect.cogstim$cognacy <- reorder(lateffect.cogstim$cognacy, lateffect.cogstim$AVERAGE)
cogstim.graph1 <- ggplot(data= lateffect.cogstim, aes(x= Stim.length, y= AVERAGE, group= cognacy,
colour= cognacy)) +
  geom_line(size= 1.5)+
  #stat_smooth(se=TRUE, method=loess)+
  geom_point(size= 5.2)+
  geom_errorbar(aes(ymax=AVERAGE+SE, ymin=AVERAGE-SE, width=0.26), size=2.4)+
  labs(y= 'Milliseconds', title= 'Latency')+
  theme(plot.title = element_text(face="bold", size=28),
```

```
      axis.text.x= element_text(size=20),
      axis.text.y= element_text(size=20),
      legend.title= element_text(size=22),
      legend.text= element_text(size=18),
      axis.title.x = element_text(size=26),
      axis.title.y = element_text(size=26))

zeffect.cogstim <- ddply(Multilink, c('cognacy', 'Stim.length'), summarize,
AVERAGE=mean(P.Latency.zm), SE=sqrt(var(P.Latency.zm)/length(P.Latency.zm)))
zeffect.cogstim$Cognacy <- reorder(zeffect.cogstim$cognacy, zeffect.cogstim$AVERAGE)
cogstim.graph2 <- ggplot(data= zeffect.cogstim, aes(x= Stim.length, y= AVERAGE, group= cognacy,
colour= cognacy)) +
  geom_line(size= 1.5)+
  #stat_smooth(se=TRUE, method=loess)+
  geom_point(size= 5.2)+
  geom_errorbar(aes(ymax=AVERAGE+SE, ymin=AVERAGE-SE, width=0.26), size=2.4)+
  labs(y= 'Milliseconds', title= 'Z-scaled')+
  theme(plot.title = element_text(face="bold", size=28),
      axis.text.x= element_text(size=20),
      axis.text.y= element_text(size=20),
      legend.title= element_text(size=22),
      legend.text= element_text(size=18),
      axis.title.x = element_text(size=26),
      axis.title.y = element_text(size=26))

png(filename='Figure 23 Latency & Z-scaled, cognacy & stimulus length effect combined.png',
width=1200, height=1066)
grid.arrange(cogstim.graph1, cogstim.graph2, nrow=2, top= textGrob('Cognacy * Stimulus Length',
gp=gpar(fontsize=34)))
dev.off()


png(filename='Figure 21 Latency, phonetic onset effect.png', width=1200, height=1066)
lateffect.on <- ddply(Multilink, c('Ph.Onset'), summarize, AVERAGE=mean(Latency),
SE=sqrt(var(Latency)/length(Latency)))
ggplot(data= lateffect.on, aes(x= Ph.Onset, y= AVERAGE, colour= Ph.Onset)) +
  #geom_line(size= 1.5)+
  geom_point(size= 6.4)+
```

```
  geom_errorbar(aes(ymax=AVERAGE+SE, ymin=AVERAGE-SE, width=0.26), size=2.4)+
  labs(y= 'Milliseconds', title= 'Phonetic Onset Interaction')+
  theme(plot.title = element_text(face="bold", size=30),
        axis.text.x= element_text(size=20),
        axis.text.y= element_text(size=20),
        legend.title= element_text(size=22),
        legend.text= element_text(size=18),
        axis.title.x = element_text(size=26),
        axis.title.y = element_text(size=26))
dev.off()

lateffect.diron <- ddply(Multilink, c('Direction', 'Ph.Onset'), summarize, AVERAGE=mean(Latency),
SE=sqrt(var(Latency)/length(Latency)))
lateffect.diron$Direction <- reorder(lateffect.diron$Direction, lateffect.diron$AVERAGE)
onset.graph1 <- ggplot(data= lateffect.diron, aes(x= Ph.Onset, y= AVERAGE, group= Direction,
colour= Direction)) +
  #geom_line(size= 1.5)+
  geom_point(size= 5.2)+
  geom_errorbar(aes(ymax=AVERAGE+SE, ymin=AVERAGE-SE, width=0.26), size=2.4)+
  labs(y= 'Milliseconds', title= 'Direction * Phonetic Onset')+
  theme(plot.title = element_text(face="bold", size=28),
        axis.text.x= element_text(size=20),
        axis.text.y= element_text(size=20),
        legend.title= element_text(size=22),
        legend.text= element_text(size=18),
        axis.title.x = element_text(size=26),
        axis.title.y = element_text(size=26))


lateffect.cogon <- ddply(Multilink, c('cognacy', 'Ph.Onset'), summarize, AVERAGE=mean(Latency),
SE=sqrt(var(Latency)/length(Latency)))
lateffect.cogon$cognacy <- reorder(lateffect.cogon$cognacy, lateffect.cogon$AVERAGE)
onset.graph2 <- ggplot(data= lateffect.cogon, aes(x= Ph.Onset, y= AVERAGE, group= cognacy,
colour= cognacy)) +
  #geom_line(size= 1.5)+
  geom_point(size= 5.2)+
  geom_errorbar(aes(ymax=AVERAGE+SE, ymin=AVERAGE-SE, width=0.26), size=2.4)+
  labs(y= 'Milliseconds', title= 'Cognacy * Phonetic Onset')+
```

```
  theme(plot.title = element_text(face="bold", size=28),
        axis.text.x= element_text(size=20),
        axis.text.y= element_text(size=20),
        legend.title= element_text(size=22),
        legend.text= element_text(size=18),
        axis.title.x = element_text(size=26),
        axis.title.y = element_text(size=26))

lateffect.freqon <- ddply(Multilink, c('freq.cat', 'Ph.Onset'), summarize, AVERAGE=mean(Latency),
SE=sqrt(var(Latency)/length(Latency)))
lateffect.freqon$freq.cat <- reorder(lateffect.freqon$freq.cat, lateffect.freqon$AVERAGE)
onset.graph3 <- ggplot(data= lateffect.freqon, aes(x= Ph.Onset, y= AVERAGE, group= freq.cat,
colour= freq.cat)) +
  #geom_line(size= 1.5)+
  geom_point(size= 5.2)+
  geom_errorbar(aes(ymax=AVERAGE+SE, ymin=AVERAGE-SE, width=0.26), size=2.4)+
  labs(y= 'Milliseconds', title= 'Frequency-category * Phonetic Onset')+
  theme(plot.title = element_text(face="bold", size=28),
        axis.text.x= element_text(size=20),
        axis.text.y= element_text(size=20),
        legend.title= element_text(size=22),
        legend.text= element_text(size=18),
        axis.title.x = element_text(size=26),
        axis.title.y = element_text(size=26))

png(filename='Figure 22 Latency, phonetic onset & direction,frequency,cognacy effect combined.png',
width=1200, height=1066)
grid.arrange(onset.graph1, onset.graph2, onset.graph3, layout_matrix = rbind(c(1,1), c(2,3)), top=
textGrob('Phonetic Onset Interactions', gp=gpar(fontsize=34)))
dev.off()


zeffect.on <- ddply(Multilink, c('Ph.Onset'), summarize, AVERAGE=mean(P.Latency.zm),
SE=sqrt(var(P.Latency.zm)/length(P.Latency.zm)))
ggplot(data= zeffect.on, aes(x= Ph.Onset, y= AVERAGE)) +
  geom_line(size= 1.5)+
  geom_point(size= 5.2)+
  geom_errorbar(aes(ymax=AVERAGE+SE, ymin=AVERAGE-SE, width=0.22), size=2.4)+
```

```
  labs(y= 'Mean Z-scaled Latency', title= 'Phonetic Onset Interaction')+
  theme(plot.title = element_text(face="bold", size=22),
        axis.title.x = element_text(size=16),
        axis.title.y = element_text(size=16))

zeffect.diron <- ddply(Multilink, c('Direction', 'Ph.Onset'), summarize,
AVERAGE=mean(P.Latency.zm), SE=sqrt(var(P.Latency.zm)/length(P.Latency.zm)))
zeffect.diron$Direction <- reorder(zeffect.diron$Direction, zeffect.diron$AVERAGE)
ggplot(data= zeffect.diron, aes(x= Ph.Onset, y= AVERAGE, group= Direction, colour= Direction)) +
  geom_line(size= 1.5)+
  geom_point(size= 5.2)+
  geom_errorbar(aes(ymax=AVERAGE+SE, ymin=AVERAGE-SE, width=0.22), size=2.4)+
  labs(y= 'Mean Z-scaled Latency', title= 'Direction * Phonetic Onset Interaction')+
  theme(plot.title = element_text(face="bold", size=22),
        axis.title.x = element_text(size=16),
        axis.title.y = element_text(size=16))

lateffect.concon <- ddply(Multilink, c('concreteness', 'Ph.Onset'), summarize,
AVERAGE=mean(Latency), SE=sqrt(var(Latency)/length(Latency)))
lateffect.concon$concretness <- reorder(lateffect.concon$Direction, lateffect.concon$AVERAGE)
ggplot(data= lateffect.concon, aes(x= concreteness, y= AVERAGE, group= Ph.Onset, colour=
Ph.Onset)) +
  #geom_line(size= 1.5)+
  stat_smooth(se=TRUE, method=loess)+
  geom_point(size= 2.0)+
  #geom_errorbar(aes(ymax=AVERAGE+SE, ymin=AVERAGE-SE, width=0.22), size=2.4)+
  labs(y= 'Mean Latency', title= 'concreteness * Phonetic Onset Interaction')+
  theme(plot.title = element_text(face="bold", size=22),
        axis.title.x = element_text(size=16),
        axis.title.y = element_text(size=16))


lateffect.dirconc <- ddply(Multilink, c('Direction', 'concreteness'), summarize,
AVERAGE=mean(Latency), SE=sqrt(var(Latency)/length(Latency)))
lateffect.dirconc$Direction <- reorder(lateffect.dirconc$Direction, lateffect.dirconc$AVERAGE)
ggplot(data= lateffect.dirconc, aes(x= concreteness, y= AVERAGE, group= Direction, colour=
Direction)) +
  #geom_line(size= 1.5)+
```

```
  stat_smooth(se=TRUE, method=loess)+
  geom_point(size= 2.0)+
  #geom_errorbar(aes(ymax=AVERAGE+SE, ymin=AVERAGE-SE, width=0.22), size=2.4)+
  labs(y= 'Mean Latency', title= 'concreteness * Phonetic Onset Interaction')+
  theme(plot.title = element_text(face="bold", size=22),
        axis.title.x = element_text(size=16),
        axis.title.y = element_text(size=16))


lateffect.stimons <- ddply(Multilink, c('Stim.length', 'Ph.Onset'), summarize,
AVERAGE=mean(Latency), SE=sqrt(var(Latency)/length(Latency)))
ggplot(data= lateffect.stimons, aes(x= Stim.length, y= AVERAGE, group= Ph.Onset, colour=
Ph.Onset)) +
  #geom_line(size= 1.5)+
  stat_smooth(se=TRUE, method=glm)+
  geom_point(size= 2.0)+
  #geom_errorbar(aes(ymax=AVERAGE+SE, ymin=AVERAGE-SE, width=0.22), size=2.4)+
  labs(y= 'Mean Latency', title= 'Stimulus Length * Phonetic Onset Interaction')+
  theme(plot.title = element_text(face="bold", size=22),
        axis.title.x = element_text(size=16),
        axis.title.y = element_text(size=16))

cyceffect.freqons <- ddply(Multilink, c('freq.cat', 'Ph.Onset'), summarize, AVERAGE=mean(Latency),
SE=sqrt(var(Latency)/length(Latency)))
ggplot(data= cyceffect.freqons, aes(x= Ph.Onset, y= AVERAGE, group= freq.cat, colour= freq.cat)) +
  geom_line(size= 1.5)+
  #stat_smooth(se=TRUE, method=glm)+
  geom_point(size= 2.0)+
  geom_errorbar(aes(ymax=AVERAGE+SE, ymin=AVERAGE-SE, width=0.22), size=2.4)+
  labs(y= 'Mean Latency', title= 'Stimulus Length * Phonetic Onset Interaction')+
  theme(plot.title = element_text(face="bold", size=22),
        axis.title.x = element_text(size=16),
        axis.title.y = element_text(size=16))



## ANOVA 2-way interaction graphs ##
# uses package 'gridExtra' and 'GGplot2'
```

```
aovinteract <-  cbind.data.frame(Multilink$Latency, Multilink$cycles,Multilink$P.Latency.lm,
Multilink$P.Latency.zm, Multilink$Direction, Multilink$cognacy, Multilink$freq.cat)
colnames(aovinteract) <- paste(c("Latency", "Cycle-time", "L-scaled", "Z-scaled", "Direction",
"Cognacy", "Frequency"), sep="")


## Latency significant effects ##
sigeffect1 <- ddply(aovinteract, c('Cognacy'), summarize, AVERAGE=mean(Latency),
SE=sqrt(var(Latency)/length(Latency)))
sigeffect1$Cognacy <- reorder(sigeffect1$Cognacy, sigeffect1$AVERAGE)
siggraph1 = ggplot(data= sigeffect1, aes(x= Cognacy, y= AVERAGE, group= Cognacy, colour=
Cognacy)) +
  #geom_line(size= 1.5)+
  geom_point(size= 5.2)+
  geom_errorbar(aes(ymax=AVERAGE+SE, ymin=AVERAGE-SE, width=0.22), size=2.4)+
  labs(y= 'Milliseconds', title= 'Cognacy')+
  theme(plot.title = element_text(face="bold", size=28),
        axis.text.x= element_text(size=17),
        axis.text.y= element_text(size=17),
        legend.title= element_text(size=18),
        legend.text= element_text(size=15),
        axis.title.x = element_text(size=25),
        axis.title.y = element_text(size=25))


sigeffect2 <- ddply(aovinteract, c('Direction'), summarize, AVERAGE=mean(Latency),
SE=sqrt(var(Latency)/length(Latency)))
sigeffect2$Direction <- reorder(sigeffect2$Direction, sigeffect2$AVERAGE)
siggraph2 = ggplot(data= sigeffect2, aes(x= Direction, y= AVERAGE, group= Direction, colour=
Direction)) +
  #geom_line(size= 1.5)+
  geom_point(size= 5.2)+
  geom_errorbar(aes(ymax=AVERAGE+SE, ymin=AVERAGE-SE, width=0.22), size=2.4)+
  labs(y= 'Milliseconds', title= 'Direction')+
  theme(plot.title = element_text(face="bold", size=28),
        axis.text.x= element_text(size=17),
        axis.text.y= element_text(size=17),
```

```
      legend.title= element_text(size=18),
      legend.text= element_text(size=15),
      axis.title.x = element_text(size=25),
      axis.title.y = element_text(size=25))

sigeffect3 <- ddply(aovinteract, c('Frequency'), summarize, AVERAGE=mean(Latency),
SE=sqrt(var(Latency)/length(Latency)))
sigeffect3$Frequency <- reorder(sigeffect3$Frequency, sigeffect3$AVERAGE)
siggraph3 = ggplot(data= sigeffect3, aes(x= Frequency, y= AVERAGE, group= Frequency, colour=
Frequency)) +
  #geom_line(size= 1.5)+
  geom_point(size= 5.2)+
  geom_errorbar(aes(ymax=AVERAGE+SE, ymin=AVERAGE-SE, width=0.22), size=2.4)+
  labs(y= 'Milliseconds', title= 'Frequency-category')+
  theme(plot.title = element_text(face="bold", size=28),
      axis.text.x= element_text(size=17),
      axis.text.y= element_text(size=17),
      legend.title= element_text(size=18),
      legend.text= element_text(size=15),
      axis.title.x = element_text(size=25),
      axis.title.y = element_text(size=25))

# marginal effects, p = 0.1
margeffect1 <- ddply(aovinteract, c('Cognacy', 'Direction'), summarize, AVERAGE=mean(Latency),
SE=sqrt(var(Latency)/length(Latency)))
margeffect1$Cognacy <- reorder(margeffect1$Cognacy, margeffect1$AVERAGE)
marggraph1 = ggplot(data= margeffect1, aes(x= Direction, y= AVERAGE, group= Cognacy, colour=
Cognacy)) +
  #geom_line(size= 1.5)+
  geom_point(size= 5.2)+
  geom_errorbar(aes(ymax=AVERAGE+SE, ymin=AVERAGE-SE, width=0.22), size=2.4)+
  labs(y= 'Milliseconds', title= 'Cognacy * Direction')+
  theme(plot.title = element_text(face="bold", size=28),
      axis.text.x= element_text(size=17),
      axis.text.y= element_text(size=17),
      legend.title= element_text(size=18),
      legend.text= element_text(size=15),
      axis.title.x = element_text(size=25),
```

```
      axis.title.y = element_text(size=25))

margeffect2 <- ddply(aovinteract, c('Cognacy', 'Frequency'), summarize, AVERAGE=mean(Latency),
SE=sqrt(var(Latency)/length(Latency)))
margeffect2$Cognacy <- reorder(margeffect2$Cognacy, margeffect2$AVERAGE)
marggraph2 = ggplot(data= margeffect2, aes(x= Frequency, y= AVERAGE, group= Cognacy, colour=
Cognacy)) +
  #geom_line(size= 1.5)+
  geom_point(size= 5.2)+
  geom_errorbar(aes(ymax=AVERAGE+SE, ymin=AVERAGE-SE, width=0.22), size=2.4)+
  labs(y= 'Milliseconds', title= 'Cognacy * Frequency-category')+
  theme(plot.title = element_text(face="bold", size=28),
      axis.text.x= element_text(size=17),
      axis.text.y= element_text(size=17),
      legend.title= element_text(size=18),
      legend.text= element_text(size=15),
      axis.title.x = element_text(size=25),
      axis.title.y = element_text(size=25))

png(filename= 'Figure xx Latency significant & marginal ANOVA interactions combined.png', width=
1200, height= 1066)
grid.arrange(siggraph1, siggraph2, siggraph3, marggraph1, marggraph2, ncol=2, top=
textGrob('Latency: Significant & Marginal Interactions', gp=gpar(fontsize=34)))
dev.off()
#layout_matrix = rbind(c(1,1,2,2,3,3), c(4,4,5))
## Cycle-time Significant interactions

cyclesig1 <- ddply(aovinteract, c('Cognacy'), summarize, AVERAGE=mean(`Cycle-time`),
SE=sqrt(var(`Cycle-time`)/length(`Cycle-time`)))
cyclesig1$Cognacy <- reorder(cyclesig1$Cognacy, cyclesig1$AVERAGE)
cyclegraph1 = ggplot(data= cyclesig1, aes(x= Cognacy, y= AVERAGE, group= Cognacy, colour=
Cognacy)) +
  geom_line(size= 1.5)+
  geom_point(size= 5.2)+
  geom_errorbar(aes(ymax=AVERAGE+SE, ymin=AVERAGE-SE, width=0.22), size=2.4)+
  labs(y= 'Mean Cycle-time', title= 'Cognacy Interaction')+
  theme(plot.title = element_text(face="bold", size=22),
      axis.title.x = element_text(size=16),
```

```
      axis.title.y = element_text(size=16))

cyclesig2 <- ddply(aovinteract, c('Frequency'), summarize, AVERAGE=mean(`Cycle-time`),
SE=sqrt(var(`Cycle-time`)/length(`Cycle-time`)))
cyclesig2$Frequency <- reorder(cyclesig2$Frequency, cyclesig2$AVERAGE)
cyclegraph2 = ggplot(data= cyclesig2, aes(x= Frequency, y= AVERAGE, group= Frequency, colour=
Frequency)) +
  geom_line(size= 1.5)+
  geom_point(size= 5.2)+
  geom_errorbar(aes(ymax=AVERAGE+SE, ymin=AVERAGE-SE, width=0.22), size=2.4)+
  labs(y= 'Mean Cycle-time', title= 'Frequency-category Interaction')+
  theme(plot.title = element_text(face="bold", size=22),
      axis.title.x = element_text(size=16),
      axis.title.y = element_text(size=16))

grid.arrange(cyclegraph1, cyclegraph2, ncol=1, top= 'Cycle-time: Significant Interactions')


## L-scaled significant interactions ##

Lscalesig1 <- ddply(aovinteract, c('Cognacy'), summarize, AVERAGE=mean(`L-scaled`),
SE=sqrt(var(`L-scaled`)/length(`L-scaled`)))
Lscalesig1$Cognacy <- reorder(Lscalesig1$Cognacy, Lscalesig1$AVERAGE)
Lscalegraph1 = ggplot(data= Lscalesig1, aes(x= Cognacy, y= AVERAGE, group= Cognacy, colour=
Cognacy)) +
  geom_line(size= 1.5)+
  geom_point(size= 5.2)+
  geom_errorbar(aes(ymax=AVERAGE+SE, ymin=AVERAGE-SE, width=0.22), size=2.4)+
  labs(y= 'Mean L-scaled Latency', title= 'Cognacy Interaction')+
  theme(plot.title = element_text(face="bold", size=22),
      axis.title.x = element_text(size=16),
      axis.title.y = element_text(size=16))

Lscalesig2 <- ddply(aovinteract, c('Frequency'), summarize, AVERAGE=mean(`L-scaled`),
SE=sqrt(var(`L-scaled`)/length(`L-scaled`)))
Lscalesig2$Frequency <- reorder(Lscalesig2$Frequency, Lscalesig2$AVERAGE)
Lscalegraph2 = ggplot(data= Lscalesig2, aes(x= Frequency, y= AVERAGE, group= Frequency,
colour= Frequency)) +
```

```
  geom_line(size= 1.5)+
  geom_point(size= 5.2)+
  geom_errorbar(aes(ymax=AVERAGE+SE, ymin=AVERAGE-SE, width=0.22), size=2.4)+
  labs(y= 'Mean L-scaled Latency', title= 'Frequency-category Interaction')+
  theme(plot.title = element_text(face="bold", size=22),
        axis.title.x = element_text(size=16),
        axis.title.y = element_text(size=16))

grid.arrange(Lscalegraph1, Lscalegraph2, ncol=1, top= 'L-scaled Latency: Significant Interactions')


## Z-scaled significant interactions ##

Zscalesig1 <- ddply(aovinteract, c('Cognacy'), summarize, AVERAGE=mean(`Z-scaled`),
SE=sqrt(var(`Z-scaled`)/length(`Z-scaled`)))
Zscalesig1$Cognacy <- reorder(Zscalesig1$Cognacy, Zscalesig1$AVERAGE)
Zscalegraph1 = ggplot(data= Zscalesig1, aes(x= Cognacy, y= AVERAGE, group= Cognacy, colour=
Cognacy)) +
  geom_line(size= 1.5)+
  geom_point(size= 5.2)+
  geom_errorbar(aes(ymax=AVERAGE+SE, ymin=AVERAGE-SE, width=0.22), size=2.4)+
  labs(y= 'Mean Z-scaled Latency', title= 'Cognacy')+
  theme(plot.title = element_text(face="bold", size=28),
        axis.text.x= element_text(size=17),
        axis.text.y= element_text(size=17),
        legend.title= element_text(size=18),
        legend.text= element_text(size=15),
        axis.title.x = element_text(size=25),
        axis.title.y = element_text(size=25))

Zscalesig2 <- ddply(aovinteract, c('Frequency'), summarize, AVERAGE=mean(`Z-scaled`),
SE=sqrt(var(`Z-scaled`)/length(`Z-scaled`)))
Zscalesig2$Frequency <- reorder(Zscalesig2$Frequency, Zscalesig2$AVERAGE)
Zscalegraph2 = ggplot(data= Zscalesig2, aes(x= Frequency, y= AVERAGE, group= Frequency,
colour= Frequency)) +
  geom_line(size= 1.5)+
  geom_point(size= 5.2)+
  geom_errorbar(aes(ymax=AVERAGE+SE, ymin=AVERAGE-SE, width=0.22), size=2.4)+
```

```
 labs(y= 'Mean Z-scaled Latency', title= 'Frequency-category')+
 theme(plot.title = element_text(face="bold", size=28),
       axis.text.x= element_text(size=17),
       axis.text.y= element_text(size=17),
       legend.title= element_text(size=18),
       legend.text= element_text(size=15),
       axis.title.x = element_text(size=25),
       axis.title.y = element_text(size=25))

png(filename= 'Figure xx Z-scaled significant ANOVA interactions combined.png', width= 1200,
height= 1066)
grid.arrange(Zscalegraph1, Zscalegraph2, ncol=1, top= textGrob('Z-scaled: Significant Interactions',
gp=gpar(fontsize=34)))
dev.off()


## Density plot ##
# uses Gplot
#largely unused in paper, but good for comparison

# 10-log frequency vs frequency categories
densityplot(~log(freq)|freq.cat, Multilink, layout=c(2,1))

# LD vs Cognacy
densityplot(~Levdist|cognacy, Multilink, layout=c(1,2))

# P.Latency.lm
densityplot(Multilink$Multilink$P.Latency.lm)

# P.Latency.zm
densityplot(Multilink$Multilink$P.Latency.zm)

# Latency
densityplot(Multilink$Latency)

# Cycles
densityplot(Multilink$cycles)
```

```
# Levenshtein Distance
densityplot(Multilink$Levdist)


## datapoint scatterplot + regression line ##
#uses package 'lattice'

# Cycle-time & Latency
print(xyplot(cycles ~ Latency, type=c("p","r"), data= Multilink))
print(xyplot(cycles ~ Latency | Direction, type=c("p","r"), data= Multilink))
print(xyplot(cycles ~ Latency | Direction:cognacy:freq.cat, type=c("p","r"), data= Multilink))

# Latency & P.Latency.lm
print(xyplot(P.Latency.lm ~ Latency, type=c("p","r"), data= Multilink))
print(xyplot(Latency ~ P.Latency.lm | Direction, type=c("p","r"), data= Multilink))
print(xyplot(Latency ~ P.Latency.lm | category, type=c("p","r"), data= Multilink))

# Latency & P.Latency.zm
print(xyplot(Latency ~ P.Latency.zm, type=c("p","r"), data= Multilink))
print(xyplot(Latency ~ P.Latency.zm | Direction, type=c("p","r"), data= Multilink))
print(xyplot(Latency ~ P.Latency.zm | category, type=c("p","r"), data= Multilink))


## Means Line-plot ##
# highlight each section to run properly
# uses Gplots
# largely unused in final paper, replaced by bar graphs below

# Latency * Cycles, Forward & Backward
par(mfrow=c(1,2), mar=c(15,5,5,5))
plotmeans(cycles~ paste(Direction), data= Multilink, las=2, xlab='')
plotmeans(Latency~ paste(Direction), data= Multilink, las=2, xlab='')

# Latency * Cycles, per condition
par(mfrow=c(1,2), mar=c(15,5,5,5))
plotmeans(cycles~ paste(Direction,cognacy,freq.cat), data= Multilink, las=2, xlab='')
plotmeans(Latency~ paste(Direction,cognacy,freq.cat), data= Multilink, las=2, xlab='')
```

```
# P.Latency.lm * Latency, Forward & Backward
par(mfrow=c(1,2), mar=c(15,5,5,5))
plotmeans(P.Latency.lm~ paste(Direction), data= Multilink, las=2, xlab='')
plotmeans(Latency~ paste(Direction), data= Multilink, las=2, xlab='')

# P.Latency.lm * Latency, per condition, highlight this section to run properly, uses Gplots
par(mfrow=c(1,2), mar=c(15,5,5,5))
plotmeans(P.Latency.lm~ paste(Direction,cognacy,freq.cat), data= Multilink, las=2, xlab='')
plotmeans(Latency~ paste(Direction,cognacy,freq.cat), data= Multilink, las=2, xlab='')

# P.Latency.zm * Latency, Forward & Backward
par(mfrow=c(1,2), mar=c(15,5,5,5))
plotmeans(P.Latency.zm~ paste(Direction), data= Multilink, las=2, xlab='')
plotmeans(Latency~ paste(Direction), data= Multilink, las=2, xlab='')

# P.Latency.zm * Latency, per condition, highlight this section to run properly, uses Gplots
par(mfrow=c(1,2), mar=c(15,5,5,5))
plotmeans(P.Latency.zm~ paste(Direction,cognacy,freq.cat), data= Multilink, las=2, xlab='')
plotmeans(Latency~ paste(Direction,cognacy,freq.cat), data= Multilink, las=2, xlab='')


## Bar graphs ##
# highlight each section to run

# Categories, Latency & P.Latency.lm
png(filename= 'Figure 14 Barchart mean latency & predicted-latency categories.png', width= 1200,
height= 1066)
par(mfrow = c(2,2))
mean.actual = tapply(Multilink$Latency,paste(Multilink$Direction,Multilink$cognacy,Multilink$freq.cat),
mean)
mean.approximate =
tapply(Multilink$P.Latency.lm,paste(Multilink$Direction,Multilink$cognacy,Multilink$freq.cat), mean)
means = cbind(mean.actual,mean.approximate)
rownames(means) <-  paste(c('1', '2', '3', '4', '5', '6', '7', '8'))
sds = cbind(tapply(Multilink$Latency,paste(Multilink$Direction,Multilink$cognacy,Multilink$freq.cat),
sd), tapply(Multilink$P.Latency.lm,paste(Multilink$Direction,Multilink$cognacy,Multilink$freq.cat), sd)
)
```

```
barplot(means,beside=T, xlab='', ylab= 'Milliseconds', names.arg=c("Latency","L-scaled Latency"),
col=c("white",'gray'), ylim=c(600,1200), xpd=F, cex.axis=2.3, cex.lab=2.0, cex.names=2.7)
title('Empirical Vs Model Comparison', cex.main=3.0, line=50)
arrows(0.5+(1:8),mean.actual,0.5+(1:8), mean.actual+sds[,1], angle=90, length=0.15)
arrows(0.5+(10:17),mean.approximate,0.5+(10:17), mean.approximate+sds[,2],
angle=90,length=0.15)
text(0.5+(1:8),670,rownames(means),srt=0, cex=2.8)
text(0.5+(10:17),670,rownames(means),srt=0, cex=2.8)

# Categories, Latency & P.Latency.zm
mean.actual = tapply(Multilink$Latency,paste(Multilink$Direction,Multilink$cognacy,Multilink$freq.cat),
mean)
mean.approximate =
tapply(Multilink$P.Latency.zm,paste(Multilink$Direction,Multilink$cognacy,Multilink$freq.cat), mean)
means = cbind(mean.actual,mean.approximate)
rownames(means) <-  paste(c('1', '2', '3', '4', '5', '6', '7', '8'))
sds = cbind(tapply(Multilink$Latency,paste(Multilink$Direction,Multilink$cognacy,Multilink$freq.cat),
sd), tapply(Multilink$P.Latency.zm,paste(Multilink$Direction,Multilink$cognacy,Multilink$freq.cat), sd)
)
barplot(means,beside=T, xlab='', ylab= 'Milliseconds', names.arg=c("Latency","Z-scaled Latency"),
col=c("white",'gray'), ylim=c(600,1200), xpd=F, cex.axis=2.3, cex.lab=2.0, cex.names=3.0)
arrows(0.5+(1:8),mean.actual,0.5+(1:8), mean.actual+sds[,1], angle=90, length=0.15)
arrows(0.5+(10:17),mean.approximate,0.5+(10:17), mean.approximate+sds[,2],
angle=90,length=0.15)
text(0.5+(1:8),670,rownames(means),srt=0, cex=2.8)
text(0.5+(10:17),670,rownames(means),srt=0, cex=2.8)

dev.off()

#Direction, P.Latency.lm
mean.actual = tapply(Multilink$Latency,paste(Multilink$Direction), mean)
mean.approximate = tapply(Multilink$P.Latency.lm,paste(Multilink$Direction), mean)
means = cbind(mean.actual,mean.approximate)
sds = cbind(tapply(Multilink$Latency,paste(Multilink$Direction), sd),
tapply(Multilink$P.Latency.lm,paste(Multilink$Direction), sd)  )
barplot(means,beside=T, xlab='', names.arg=c("Latency","L-scaled Latency"), col=c("white",'gray'),
ylim=c(600,1000), xpd=F)
arrows(0.5+(1:2),mean.actual,0.5+(1:2), mean.actual+sds[,1], angle=90, length=0.1)
arrows(1.5+(3:4),mean.approximate,1.5+(3:4), mean.approximate+sds[,2], angle=90,length=0.1)
```

```
text(0.5+(1:2),400,rownames(means),srt=90)
text(1.5+(3:4),400,rownames(means),srt=90)



#Direction, Latency & P.Latency.zm
mean.actual = tapply(Multilink$Latency,paste(Multilink$Direction), mean)
mean.approximate = tapply(Multilink$P.Latency.zm,paste(Multilink$Direction), mean)
means = cbind(mean.actual,mean.approximate)
sds = cbind(tapply(Multilink$Latency,paste(Multilink$Direction), sd),
tapply(Multilink$P.Latency.lm,paste(Multilink$Direction), sd)  )
barplot(means,beside=T, xlab='', names.arg=c("Latency","Z-scaled Latency"), col=c("white",'gray'),
ylim=c(600,1000), xpd=F)
arrows(0.5+(1:2),mean.actual,0.5+(1:2), mean.actual+sds[,1], angle=90, length=0.1)
arrows(1.5+(3:4),mean.approximate,1.5+(3:4), mean.approximate+sds[,2], angle=90,length=0.1)
text(0.5+(1:2),400,rownames(means),srt=90)
text(1.5+(3:4),400,rownames(means),srt=90)



# Direction, Cycles
png(filename= 'Figure 13 Barchart mean latency & cycle-time, direction.png', width= 1200, height=
1066)
par(mfrow = c(2,2))
mean.cycles = tapply(Multilink$cycles,paste(Multilink$Direction), mean)
means = cbind(mean.cycles)
sds = cbind(tapply(Multilink$cycles,paste(Multilink$Direction), sd) )
barplot(means,beside=T, xlab='', ylab='Cycles', names.arg=c('Cycle-time'), col=c("white",'gray'),
ylim=c(15,30), xpd=F, cex.axis=2.3, cex.lab=2.0, cex.names=3.0)
arrows(0.5+(1:2),mean.cycles,0.5+(1:2), mean.cycles+sds[,1], angle=90, length=0.15)
text(0.5+(1:2),19,rownames(means),srt=0, cex=4.4)

#Direction, Latency
mean.Latency = tapply(Multilink$Latency,paste(Multilink$Direction), mean)
means = cbind(mean.Latency)
sds = cbind(tapply(Multilink$Latency,paste(Multilink$Direction), sd),
tapply(Multilink$Latency,paste(Multilink$Direction), sd)  )
barplot(means,beside=T, xlab='', ylab='Milliseconds', names.arg=c('Latency'), col=c("white",'gray'),
ylim=c(600,1000), xpd=F, cex.axis=2.3, cex.lab=2.0, cex.names=3.0)
```

```
arrows(0.5+(1:2),mean.Latency,0.5+(1:2), mean.Latency+sds[,1], angle=90, length=0.15)
text(0.5+(1:2),700,rownames(means),srt=0, cex=4.4)

dev.off()


# Categories, Cycles
mean.cycles = tapply(Multilink$cycles,paste(Multilink$Direction,Multilink$cognacy,Multilink$freq.cat),
mean)
means = cbind(mean.cycles)
sds = cbind(tapply(Multilink$cycles,paste(Multilink$Direction,Multilink$cognacy,Multilink$freq.cat), sd)
)
barplot(means,beside=T, xlab='', names.arg=c('Cycle-time'), col=c("white",'gray'), ylim=c(15,30),
xpd=F)
arrows(0.5+(1:8),mean.cycles,0.5+(1:8), mean.cycles+sds[,1], angle=90, length=0.1)
text(0.5+(1:8),10,rownames(means),srt=90)

#Categories, Latency
mean.Latency =
tapply(Multilink$Latency,paste(Multilink$Direction,Multilink$cognacy,Multilink$freq.cat), mean)
means = cbind(mean.Latency)
sds = cbind(tapply(Multilink$Latency,paste(Multilink$Direction,Multilink$cognacy,Multilink$freq.cat),
sd), tapply(Multilink$Latency,paste(Multilink$Direction,Multilink$cognacy,Multilink$freq.cat), sd)  )
barplot(means,beside=T, xlab='', names.arg=c('Latency'), col=c("white",'gray'), ylim=c(600,1200),
xpd=F)
arrows(0.5+(1:8),mean.Latency,0.5+(1:8), mean.Latency+sds[,1], angle=90, length=0.1)
text(0.5+(1:8),350,rownames(means),srt=90)


##########################################################################
# thank you to my favourite bands for providing music to code this:
# Matt Pond PA, Lights, Forgive Durden, Land Of Talk, The Academy Is..., and Coeur De Pirate
# life is empty without your art, and I couldn't have done this without you.
##########################################################################

write.table(Multilink, file= 'Multilink2.csv', quote=FALSE, sep=',', dec= '.', row.names=FALSE,
col.names=TRUE)
```

Thank you for reading,

Bedankt voor het lezen,

Merci d'avoir lu,

Danke für lesen,

Sincerely,

Jesse Peacock